Assays for lycosyltransferases

Goos, Niina

Awarding institution:
King's College London

Assays for Glycosyltransferases

A thesis submitted for the degree of Doctor of Philosophy of King's College London

Niina Göös

Institute of Pharmaceutical Science,
King's College London, Franklin-Wilkins Building,
150 Stamford Street, London, SE1 9NH.
August, 2013

Abstract

Glycosyltransferases (GTs) are a large enzyme family that are involved in the biosynthesis of complex carbohydrates and glycoconjugates. These glycosylation reactions are essential to many fundamental biological processes, including cellular adhesion, cell signalling and bacterial cell wall biosynthesis. The aim of this project was the development of operationally simple assays for GTs and their application for the identification and characterisation of novel GT inhibitors. The main focus was on the bacterial GT, α1,4-galactosyltransferase C (LgtC) which transfers galactose from uridine-diphosphate-galactose to the terminal lactose on the cell wall lipooligopolysacharide of *Neisseria meningitidis*.

A fluorescence-based ligand displacement assay (LDA) was optimised and validated for compound library screening. Additionally, the applicability of the LDA was investigated with three other GTs: bovine β1,4-galactosyltransferase, TcdB from *Clostridium difficile* and NGT from *Actinobacillus pleuropneumoniae*. The validated LDA was used for screening two compound libraries (c. 400 compounds in total) against LgtC for the first time. Positive hits were identified from both screening campaigns and selected compounds from two structural classes were characterised. One of these inhibitor classes was identified as molecular aggregators, which inhibit LgtC non-specifically. Results from the enzymological characterisation of the second class of hit compounds suggest that these inhibitors were mixed-type inhibitors and inhibit LgtC in micromolar range with considerable potential for further development.

Development of a novel biochemical assay was attempted, in order to characterise potential inhibitors from the compound screening. This assay was based on an unnatural fluorescent acceptor and coupling of the GT reaction to a glycosidase reaction. During the development of this novel biochemical assay, unexpected enzymological features of LgtC were discovered. LC-MS/MS was used to investigate unknown products originating from the LgtC-catalysed reaction. The results suggest LgtC transfers more than one sugar moiety to the unnatural acceptor. Additionally two existing biochemical assays: HPLC-based assay and phosphatase coupled assay were successfully optimised and validated for determining the kinetic parameters of the LgtC catalysed reaction.

Table of Contents

Abstract **2**

List of figures **7**

List of schemes **14**

List of tables **15**

Abbreviations **18**

Acknowledgements **23**

1. Introduction **24**

1.1. GLYCOSYLTRANSFERASES 25

1.1.1. Classification of glycosyltransferases *27*

1.1.2. Mechanism of GTs *28*

1.1.3. Glycosyltransferases in biological systems *30*

1.1.4. Inhibitors of glycosyltransferases *33*

1.2. GLYCOSYLTRANSFERASE ASSAYS 35

1.2.1. Non fluorescence-based assays *38*

1.2.2. Fluorescence-based assays *40*

1.3. GLYCOSIDASES 43

1.4. PROJECT OBJECTIVES 44

2. Fluorescence-based ligand displacement assay **46**

2.1. INTRODUCTION 46

2.1.1. LgtC *47*

2.1.2. β1,4-GalT *49*

2.1.3. TcdB *51*

2.1.4. NGT *52*

2.2. FLUORESCENCE-BASED DISPLACEMENT ASSAY 53

2.2.1. Investigating: reaction volume, concentration of fluorophore and DMSO, buffer and incubation time *55*

2.2.2. Influence of LgtC activation *58*

2.2.3. Influence of metal *60*

2.2.4. Non-specific binding and influence of surfactant *62*

2.2.5. Influence of LgtC concentration 65

2.2.6. Summary of ligand displacement assay optimisation and determination of assay reproducibility 66

2.3. STABILITY OF LGTC 68

2.4. ASSAY ADAPTATION TO 384 WELL MICROPLATES 70

2.4.1. Linearity of fluorophore and reaction volume 70

2.4.2. Fluorescence based ligand displacement assay 71

2.4.3. Ligand displacement with other sugar nucleotides and nucleotides 72

2.5 INVESTIGATION OF OTHER GLYCOSYLTRANSFERASES 74

2.6 SUMMARY AND CONCLUSIONS 78

2.7 EXPERIMENTAL 81

2.7.1 Expression and purification of LgtC 81

2.7.2 Concentration and activity of LgtC 83

2.7.3 General settings 83

2.7.4 Assay parameters; reaction volume, concentration of fluorophore and DMSO, buffer and incubation time 85

2.7.5 Activation of LgtC 86

2.7.6 Influence of metal 86

2.7.7 Non-specific binding and influence of surfactant 86

2.7.8 Influence of LgtC concentration 87

2.7.9 Reproducibility experiment 87

2.7.10 Stability of LgtC 88

2.7.11 Assay adaptation to 384 well microplates 88

2.7.12 Investigation of other GTs 89

3 Biochemical assays 90

3.1 INTRODUCTION 90

3.2 HPLC METHOD 91

3.2.1 Method optimisation 92

3.2.2 System suitability 94

3.2.3 Adaptation of the HPLC method for monitoring LgtC catalysed reaction 97

3.3 DEVELOPMENT OF GLYCOSIDASE COUPLED ASSAY 100

3.3.1 Principle of the glycosidase coupled assay 101

3.3.2 Optimising β-galactosidase catalysed reaction 103

3.3.3 Coupling LgtC reaction and β-galactosidase reactions 105

3.3.4 LCMS method development for the identification of the unknown product..... 109
3.3.4.1 Underivatised HILIC Silica bonded column 111
3.3.4.2 ZIC-pHILIC column..... 113
3.3.4.3 Amide bonded HILIC column 116
3.3.5 Optimisation of LgtC catalysed reaction conditions for LC-MS..... 120
3.3.6 Analysis of the reaction product by LC-ESI-MS/MS..... 121
3.4 PHOSPHATASE COUPLED BIOCHEMICAL ASSAY..... 125
3.4.1 Principle of the assay 125
3.4.2 Optimising the assay conditions with LgtC 126
3.5 SUMMARY AND CONCLUSIONS..... 130
3.6 EXPERIMENTAL..... 134
3.6.1 HPLC 134
3.6.2 Glycosidase coupled assay 135
3.6.2.1 Optimising the β-galactosidase concentration, calibration curve and stability of resorufin 136
3.6.2.2 Coupling LgtC and β-galactosidase reactions 136
3.6.2.3 LCMS method development 138
3.6.3 Phosphatase coupled assay 141
3.6.3.1 Optimisation with LgtC 141
4 Compound library screening and compound characterisation 144
4.1 INTRODUCTION 144
4.1.1 Compound library screening protocol..... 146
4.2 DIVERSE COMPOUND LIBRARY SCREENING..... 150
4.2.1 Characterisation of the hits from the diverse compound library..... 152
4.2.1.1 Synthesis of pyrimidopyrimidine compounds 154
4.2.1.2 Determination of IC_{50} by ligand displacement assay 157
4.2.1.3 Determination of K_i by phosphatase coupled assay and investigation of aggregation-based inhibition..... 158
4.2.1.4 Determination of particle size by DLS..... 162
4.3 TARGET-FOCUSED LIBRARY SCREENING 164
4.3.1 Enzymological characterisation of hits from the focus-target library 167
4.3.1.1 Determination of K_i..... 168
4.3.1.2 Determination of IC_{50}..... 171
4.3.1.3 Determination of inhibition mode of pyrazolones 172
4.4 SUMMARY AND CONCLUSIONS..... 175

4.5 EXPERIMENTAL 178
4.5.1 Diverse compound library screening 178
4.5.2 Target-focused library screening 179
4.5.3 The hit characterisation studies 181
4.5.3.1 Determination of IC_{50} and pre-incubation study by the LDA 182
4.5.3.2 Determination of percentage inhibition by HPLC-based assay 182
4.5.3.3 Determination of activity of phosphatase in the presence of inhibitor 183
4.5.3.4 Determination of K_i by PCA 183
4.5.3.5 Determination of mode of inhibition by PCA 183
4.5.4 Synthesis of PP compounds 184
4.5.5 Determination of λ_{max} and purity of the synthesised PP compounds 191
4.5.6 Determination of particle size by DLS 192
5 Summary and future work 193
Appendix 196
1. PHOSPHATASE COUPLED ASSAY DEVELOPMENT WITH B1,4-GALT 196
2. LIST OF REAGENTS 200
References 202

List of figures

Figure 1.1. Glycosyltransferase catalysed reaction 25

Figure 1.2. Overall folds observed for GTs. The GT-A fold is represented by the inverting GT SpsA from *Bacillus subtilis* in the presence of UDP (Protein Data Bank, PDB: 1QGQ)[20] and the GT-B fold, by bacteriophage T4 β-glucosyltransferase in the presence of UDP (PDB: 1JG7)[23]. α-helixes are in blue, β-strands in red and donor substrates are presented in stick form (yellow)..... 28

Figure 1.3. Retaining and inverting transfer reactions catalysed by glycosyltransferases[5]..... 29

Figure 1.4. Direct-displacement mechanisms of inverting GTs (A) and a proposed double-displacement mechanism for retaining GTs (B). R_1 would generally be a nucleoside mono- or diphosphate of the donor substrate and R_2 would correspond to the acceptor substrate 29

Figure 1.5. Structure of two cancer-associated glycans: sialyl Lewis X and Lewis Y[33] 31

Figure 1.6. Schematic presentation of the cell wall of Gram-negative bacteria (modified from reference[42]) 32

Figure 1.7. LOS of *N. meningitidis*. LgtC transfers galactose to the terminal end of the α-chain if LgtA is inactive (the figure is modified from reference[43])..... 33

Figure 1.8. A) UDP-carba-Glc, a donor substrate analogue B) An acceptor analogue, 2-naphthyl 2-butanamido-2-deoxy-1-β-D-glucopyranoside inhibitor of β1,4-GalT[61], C) An acceptor analogue, per-O-acetylated GlcNAc-β1,3-Gal-β-*O*-naphtalenemethanol, inhibitor of β1.4-GalT T1[62], D-F) small molecular inhibitors of *O*-GlcNAc transferase identified by HTS[56].... 34

Figure 1.9. Chemical structure of Miglustat[65] 35

Figure 1.10. Fluorescently labelled donor analogues used in ligand displacement assays for inhibitor screening against Murg by Walker's group[89] (A) and against UDP-galactopyranose mutase by Kiessling's group[91] (B) 42

Figure 1.11. Chromogenic substrates for the analysis of α-galactosidase. A) 4-methylumbelliferyl-α-D-galactopyranoside, B) resorufinyl-α-D-galactopyranoside and C) *p*-nitrophenyl-α-D-galactopyranoside..... 43

Figure 1.12. Fluorescent UDP-Gal analogue (5-(5-formyl-2-thiophene)-UDP-Gal) and principle of the ligand displacement assay 45

Figure 2.1. Crystal structure of LgtC complex with UDP-2FGal in yellow (ball- and- stick model) Flexible loops are coloured in green, β-sheets in red and α-helixes are in blue (PDB: 1G9R) 48

Figure 2.2. A significant structural change in open (Green, conf. I) and closed (Magenta, conf. II) conformations of β1,4-GalT. The long loop (residues I345-H365) and the short loop containing W314 (ball-and-stick model). W314 is facing in to the catalytic pocket in closed conformation (active conformation). (Taken from reference[121]) 50
Figure 2.3. Structure of the fluorophore used in the LDA, 5-(5-formyl-2-thiophene)-UDP-Gal. 53
Figure 2.4. Principle of the ligand displacement assay. A) Fluorescence intensity quenches when the ligand (5FTUDP-Gal) is incubated with increasing concentration of a protein (*e.g.* LgtC). B) Fluorescence intensity is restored with increasing concentration of a high affinity binder or inhibitor (*e.g.* UDP) 54
Figure 2.5. Linearity of fluorophore in the range of 39 nM to 2500 nM at total volume of 100 µL and 200 µL..... 55
Figure 2.6. Ligand displacement assay with various UDP-Gal concentrations in Tris/ HCl (A) and in HEPES (B) buffers in various pH 56
Figure 2.7. Displacement assay with UDP-Gal (A) and UDP (B) with various incubation times . 57
Figure 2.8. Effect of DMSO on the LDA with various concentrations of UDP-Gal (A) and UDP (B) 58
Figure 2.9. Effect of LgtC activation with DTT compared with non-activated LgtC in the binding affinity assay 59
Figure 2.10. Effect of manganese for fluorophore binding to LgtC. Fluorophore binding affinity monitored with increasing concentration of LgtC without Mn^{2+} addition and in the presence of 10 mM of Mn^{2+}..... 60
Figure 2.11. Effect of manganese ion (10 mM) in the ligand displacement assay with UDP-Gal (A) and UDP (B) 61
Figure 2.12. Investigation of non-specific binding with BSA as a control. Binding affinity of the fluorophore with increasing concentration of LgtC or BSA. Steps of quenching marked as 1-3 on the blue curve 62
Figure 2.13. Fluorophore binding to LgtC in the presence of various TX-100 concentrations ... 63
Figure 2.14. Influence on TX-100 at 0, 0.001 and 0.01% for fluorophore displacement with UDP and IC_{50} of UDP, respectively 64
Figure 2.15. Fluorophore displacement with UDP at 1.0, 5.0 and 15 µM of LgtC, and IC_{50} of UDP respectively. The LDA performed without the addition of Mn^{2+} and in the presence of TX-100 (0.01 %) 65

Figure 2.16. Change in binding affinity (A) and the fluorophore displacement (B) after 2 and 5 months expression of LgtC. IC_{50} of UDP-Gal decreased from 32 µM to 12 µM over the time 69
Figure 2.17. Comparison of 96 and 384 well microplates. Binding of fluorophore increasing concentration of LgtC (A) and ligand displacement with UDP (B) determined by the optimised assay parameters (Chapter 2.2.6). IC_{50} of UDP 28 ± 4 µM and 22 ± 1 µM in 384 and 96 well microplates, respectively..... 71
Figure 2.18. Fluorophore displacement with UDP-Gal, UDP, UMP and uridine determined with the optimised LDA (DMSO (10 %), TX-100 (0.01 %) in 13 mM HEPES (pH 7.0) buffer) 72
Figure 2.19. Structure of fluorescent UDP-Glc analogue (5-(5-formyl-2-thiophene)-UDP-Glc, 5FTUDP-Glc) used studying NGT and TcdB 74
Figure 2.20. Fluorophore binding monitored with increasing concentrations of LgtC, TcdB, β1.4-GalT and NGT. The experiments performed in the presence of TX-100 and DMSO 75
Figure 2.21. Molecular surface presentation of the active sites of A) LgtC (PDB: 1G9R), B) β1,4-GalT (PDB: 1FR8), C) NGT (PDB: 3Q3H), D) TcdB (PDB: 2BVL), and E) α1,3-GalT (PDB: 1K4V) in the presence of donor substrates. Donor substrates (red) are shown in stick form and loop regions are shown in green..... 76
Figure 2.22. SDS-PAGE analysis of collected fractions. Lane 1 is a protein ladder. Lanes 2-3 are flow-through, lanes 4-6 are washes and 7-9 are collected LgtC fractions. Lane 7 is the first collected fraction (concentration not determined), lane 10 is LgtC from previous batch. Concentration of LgtC on lanes 8-10 are 11.84, 5.6 and 19.55 mg/mL respectively..... 82
Figure 2.23. General microplate layout in the binding affinity experiments (96 well microplate). Enzyme (E) and mixture containing fluorophore, DMSO and TX-100 (M). All dilutions in 13 mM HEPES buffer (pH 7.0) 84
Figure 2.24. General microplate layout in the ligand displacement (IC_{50}) experiments (96 well microplate). Enzyme (E), binder (I) and mixture containing fluorophore and TX-100 (M). All dilutions in 13 mM HEPES buffer (pH 7.0) 85
Figure 2.25. Microplate layout in the reproducibility experiment. Negative control (N), positive control (P) and control (C)..... 87
Figure 3.1. UMP, UDP-Gal and UDP eluted at 5.8, 9.4 and 12.8 min, respectively. Separations were performed with the optimised gradient (Table 3.1), an RP column and the wavelength of UV detection was 265 nm 94
Figure 3.2. Peak area increased linearly with increasing concentration of UDP-Gal and UDP in range of 1.0 µM to 200 µM with optimised HPLC method..... 96

Figure 3.3. UDP formation of non-activated versus activated LgtC catalysed reaction followed by HPLC-based method with detection wavelength 265 nm. Lactose and UDP-Gal concentrations 2.0 mM and 200 µM, respectively 98
Figure 3.4. Hydrolysis of Rgal (20 µM) with various β-galactosidase concentrations (A). Fluorescence linearity of resorufin by hydrolysis of 0.8 to 50 µM Rgal by 0.5 U/mL of β-galactosidase (B). Stability of resorufin at concentrations of 0.4 to 50 µM (C). Fluorescence monitored at λ_{em} =590 nm and λ_{ex} =544 nm 104
Figure 3.5. Reaction process of LgtC catalysed reaction monitored by optimised HPLC method (Chapter 3.2.1). The reaction monitored at 0.005 U/mL and at 0.05 U/mL of LgtC, and compared to the peak area of UDP-Gal in the negative control where LgtC was not added. Rgal at 100 µM was used as an acceptor and UDP-Gal at 60 µM as a donor. 105
Figure 3.6. Negative and positive control investigations by coupling LgtC and β-galactosidase reactions. The experiment was performed at fixed concentrations of UDP-Gal (50 µM), Rgal (6.3 µM), LgtC (0.1 U/mL) and β-galactosidase (0.5 U/mL). Fluorescence monitored at λ_{em} =590 nm and λ_{ex} =544 nm 106
Figure 3.7. Glycosidase coupled reaction in various UDP-Gal concentrations and various incubation times. Reactions were carried out at 6.3 µM of Rgal, 0.1 U/mL LgtC and 0.5 U/mL β-galactosidase. Fluorescence monitored at λ_{em} =590 nm and λ_{ex} =544 nm 107
Figure 3.8. LgtC catalysed reaction stopped after 30 minutes incubation (blue trace) overlaid with the control, in the absence of LgtC (red trace) (A). LgtC catalysed reaction progress monitored by HPLC-based method (B). The reaction was carried out at 200 µM of UDP-Gal and Rgal, 0.02 U/mL of LgtC in 50 mM HEPES buffer including 1.0 mM $MnCl_2$ and $MgCl_2$ (pH 7.1). All components analysed at 265 nm 108
Figure 3.9. Different types of silanol groups on the silica surface: A) single, B) geminal and C) vicinal 112
Figure 3.10. Separation of mixture of UDP, UDP-Gal and Rgal (100 µg/mL each) with bare silica column. UDP-Gal and UDP was detected at 265 nm (blue trace), and Rgal at 470 nm (red trace) 113
Figure 3.11. Sulfoalkylbetaine zwitterionic stationary phase 114
Figure 3.12. Mixture of UDP, UDP-Gal and Rgal (50 µg/mL each) analysed with ZIC-*p*HILIC column at 30 °C (blue trace) and 60 °C (red trace), detection wavelength 265 nm 114
Figure 3.13. Amide modified silica stationary phase 116

Figure 3.14. A) Mixture of Rgal, UDP-Gal and UDP (green trace), 50 µg/mL each B) UDP (red trace) and UDP-Gal (blue trace), 100 µg/mL each examined with two separate HPLC runs. Detection wavelength 265 nm 117
Figure 3.15. Separation of the mixture of Rgal, UDP and UDP-Gal (50 µg/mL each). Overlay of HPLC runs in A) gradient 0-100 % of 10 mM NH_4Ac:ACN (50:50) over 15 minutes against buffer A (10 mM NH_4Ac:ACN (5:95)) B) gradient 0-100 % of 10 mM NH_4Ac:ACN (70:30) over 15 minutes against buffer A and C) the same as trace B but gradient over 20 minutes. Detection wavelength 265 nm 118
Figure 3.16. Influence of various water content in the sample solvent, 10 %, 20 % and 30 % of water in green, red and blue traces, respectively. A) with Rgal, detected at 470 nm and B) with UMP, UDP-Gal and UDP (respectively, in elution order) were detected at 265 nm 119
Figure 3.17. HPLC chromatogram of LgtC catalysed reaction, Rgal as an acceptor. A new peak appeared at 13 minutes (blue trace), blank and control traces in green and red, respectively. Compounds separated with Amide HILIC column and detected at 265 nm 120
Figure 3.18. LC-ESI-MS/MS analysis of the standards of UMP, UDP, UDP-Gal and Rgal. The *m/z* fragmentation of the analysed peaks presented on next to the peak for each standard 122
Figure 3.19. Glycosylated products of LgtC catalysed reaction: Rgal-gal (8.5 minutes) and Rgal-gal-gal (10.2 minutes), and the corresponding masses below, respectively 123
Figure 3.20. LgtC catalysed reaction in the presence and in the absence of CEL 127
Figure 3.21. UDP-Gal hydrolysis in the presence and in the absence of LgtC at 20 nM and 50 nM 128
Figure 3.22. Formation of phosphate with increasing $MnCl_2$ concentration and at fixed concentrations of UDP-Gal and LgtC 128
Figure 3.23. Colour development of the malachite reagents at fixed phosphatase concentration and in various TX-100 and UDP concentrations 129
Figure 3.24. Layout of the micro plate in K_m experiments. Columns 2-3: UDP calibration (phosphatase control), rows B-D from column 4: various UDP-Gal in the presence of LgtC, rows E-G from column 4: background at each UDP-Gal concentration in the absence of LgtC. Master mixture (M) contain lactose, $MnCl_2$, CEL and CIP, UDP at concentrations of 0, 0.78, 1.56, 3.13, 6.25 and 12.5 µM (U1-U5), UDP-Gal (G) and LgtC (E) 143
Figure 4.1. Tested thiazolidinones at 20 µM and 100 µM in the presence of the fluorophore. The control is pure fluorophore and UDP at 5.0 mM in the presence of the fluorophore. Fluorescence intensity was measured at λ_{ex}= 350 nm and λ_{em}= 430 nm which are relevant to the fluorophore, 5FTUDP-Gal 148

Figure 4.2. Fluorescence intensity (FI) comparison between the controls and T1 and T3 at 100 µM in the absence and in the presence of LgtC. I is binder: T1, T3 or UDP, E is LgtC and F is the fluorophore (5FTUDP-Gal). Δ1 is the assay window: the maximum FI change between negative and positive control (UDP 5.0 mM), Δ2 and Δ4 is the difference in FI caused by the tested compounds compared to the control (F only), Δ3 and Δ5 is the difference in FI between the corrected negative control (corrected neg) and the FI signal of the test compounds in the presence of F 149
Figure 4.3. Examples of compounds included in the DOS library[217] 151
Figure 4.4. The general scaffold of the positive hits from the primary screen of the DOS library 151
Figure 4.5. Fluorophore displacement with one of the hit compounds from the primary screen (P2_27) in various concentrations against LgtC. Experiment carried out in the presence of 10 % DMSO and fluorophore displacement monitored at λ_{ex}= 350 nm and λ_{em}= 430 nm 152
Figure 4.6. Determination of IC_{50} for 1c in the presence and in the absence of TX-100 by the LDA 158
Figure 4.7. Structure of benzoyl pyrimidine, BP 159
Figure 4.8. Phosphatase control by the PCA. The functionality of the phosphatase was not affected by UDP formation in various concentrations of 1c or P2_93. Experiment was carried out at fixed concentrations of UDP (6.25 µM) and phosphatase (10 U/mL)..... 160
Figure 4.9. Particle size of 1b at 10 µM, 50 µM and 100 µM. Optimum concentration for DLS experiments was found between 10 µ and 50 µM of 1b..... 163
Figure 4.10. Percentage inhibition/fluorophore displacement at 25 µM of the compound against LgtC obtained by the LDA and the PCA 166
Figure 4.11. Non-competitive/ mixed inhibition. Inhibitor (I) binds at a separate site and may bind either to E (enzyme) or/and to ES (enzyme-substrate) complex (on the left). Lineweaver-Burk plot of mixed inhibition (on the right) 167
Figure 4.12. Phosphatase control by the PCA. The functionality of the phosphatase was not affected by UDP formation in various concentrations of p5 or p18. Experiment was carried out at fixed concentrations of UDP (6.25 µM) and phosphatase (10 U/mL)..... 169
Figure 4.13. Dixon plot of reciprocal initial velocity (v) versus p5 concentration to calculate the K_i value. K_i determination against LgtC in the absence and in the presence of TX-100, and against β1,4-GalT by the PCA..... 169
Figure 4.14. IC_{50} and pre-incubation study with p5 and p18 by the LDA in the presence of TX-100. A) Incubation 1, B) Incubation 2, C) Incubation 3, D) Incubation 3, the displacement

of the fluorophore monitored in various concentration of p18 over 30 minutes after p18 was added 171
Figure 4.15. Michaelis-Menten and Lineweaver-Burk plot in the absence and in the presence of p5 (A and B) or p18 (C and D) 173
Figure 4.16. Cornish-Bowden plot (A and C) and Dixon plot (B and D) in various p5 and p18 concentrations. The intersection point on both plots suggests p5 and p18 are mixed mode inhibitors 174
Figure 4.17. Z' factor reduced during the compound screening. Micro plate number 1 was prepared first and so forth to the last microplate number 9 176
Figure 4.18. Layout of the microplate on the compound screening 179
Figure 4.19. Layout of the microplate in compound screening by PCA. Columns 2-3; UDP calibration, columns 4-5; various compounds and negative control on row B, columns 6-7; background for each compound. Master mixture (M), UDP in various concentrations, 0, 0.78, 1.56, 3.13, 6.25 and 12.5 µM (U1-U5), UDP-Gal (G), LgtC (E) and various compounds (I) 181
Figure 5.1. A) Possible outcome of the LDA when the fluorophore is displaced with pyrazolones. B) The principle of the fluorophore displacement in the presence of acceptor in order to investigate if pyrazolones bind to the acceptor binding site. Abbreviations: A; acceptor, F; fluorophore, I; inhibitor 195

List of schemes

Scheme 1.1. Principles of assays used for determination of GT activity and ligand binding. A) Native GT reaction; NDP can be detected by HPLC/UV, coupled enzymatic assay, immunodetection or chemosensor and primary product can be detected by MS, B) Labelled donor or acceptor: the product can be detected by spectrophotometric or radiochemical methods, C) FRET based assay principle, D) fluorescence-based ligand displacement assay for detection GT inhibitors (modified from reference[63]) 37

Scheme 2.1. LgtC catalysed reaction 48

Scheme 3.1. LgtC catalysed reaction can be monitored by HPLC/UV due to the UV-active catalytic base of UDP 92

Scheme 3.2. General principle of the glycosidase coupled assay. LgtC catalyses the transfer of galactose from UDP-Gal to the fluorogenic acceptor producing glycosylated product which cannot be hydrolysed by specific glycosidase and fluorescence remains low. The unreacted acceptor is then hydrolysed by glycosidase resulting in high fluorescence 101

Scheme 3.3. Resorufin-β-galactopyranoside (Rgal) hydrolysed by β-galactosidase results strongly in fluorescent resorufin (λ_{em} =590 nm, λ_{ex} =570 nm) 102

Scheme 3.4. LgtC catalysed reaction when Rgal used as an acceptor 110

Scheme 3.5. Phosphatase coupled GT assay can be applied to any GT reaction where the leaving group contain a removable phosphate (*i.e.* UDP from UDP-Gal). (The scheme is modified from reference[85]) 125

Scheme 4.1. Schematic presentation of design of focused compound libraries. Uridine analogues (1-3) with alternative linkers mimic the structure of natural donor. Various functional groups (R) replace the natural sugar. Human 4-epimerase inhibitor discovered from the uridine analogue library.[215] 144

Scheme 4.2. Synthesis of PP analogues 1b to 1g via Biginelli-type reaction. The functional roups (R_1 and R_2) and yields are presented in Table 4.3 155

Scheme 4.3. Synthesis of PP analogues 2b to 2c. The functional roups (R_1 and R_2) and yields are presented in Table 4.3 155

List of tables

Table 1.1. Nucleotide sugar donors in mammalian metabolism 26
Table 2.1. IC_{50} for UDP-Gal in Tris/HCl and HEPES buffers at various pH 56
Table 2.2. IC_{50} values in different incubation modes and in various final incubation times. Experiments performed at 30 °C 57
Table 2.3. Comparison of the "original" and the optimised LDA parameters 66
Table 2.4. Concentrations of various LgtC batches and determined IC_{50} values with the optimised LDA 67
Table 2.5. Z' factor and S/B ratio of the optimised LDA 67
Table 2.6. Change in specific activity, K_m, K_{cat} and IC_{50} of UDP-Gal determined with LgtC after 2 and 5 months after the expression 69
Table 2.7. Linearity of fluorophore in range of 12.5 to 1000 nM (10 measurement points) in various reaction volumes performed in the assay buffer (13 mM HEPES pH 7.0) 71
Table 2.8. IC_{50} values for various nucleotides determined the optimised LDA and the published results of the same experiment (both assay conditions described in detail in Chapter 2.2.6) 73
Table 2.9. Properties of studied GTs 74
Table 2.10. Summary of all GTs investigated with the binding affinity experiments in the current and in the previous study 80
Table 2.11. The compositions of loading, washing and eluting buffers used in LgtC purification 81
Table 2.12. Concentrations of the assay components in the binding affinity and ligand displacement assays 84
Table 2.13. Conditions used in binding affinity assay for studied GTs 89
Table 3.1. Gradient steps of HPLC method for analysis of UMP, UDP-Gal and UDP 93
Table 3.2. System suitability tests determined at 25 µM of UMP, UDP-Gal and UDP. %RSD is in brackets 96
Table 3.3. Reproducibility of LgtC catalysed reaction by monitoring UDP peak area (*n*= 8) 98
Table 3.4. K_m and K_{cat} of UDP-Gal determined by the optimised HPLC-based method (*n*= 3) 99
Table 3.5. cLogP values of the analytes. Determined by ACD/3D Viewer Freeware 12.01 111
Table 3.6. Measured parameters of UDP, UDP-Gal and Rgal (50 µg/mL each) analysed with various gradients at 60 °C. Resolution (R), tailing factor (T) and retention time (RT) 115

Table 3.7. Reproducibility of retention time (RT) and resolution (R) for Rgal, UMP, UDP-Gal and UDP (50 µg/mL each), (*n*=6). Experiment performed with HPLC gradient: 0-100 % of buffer B (10 mM NH_4Ac:ACN (70:30)) over 20 minutes against buffer A (10 mM NH_4Ac:ACN (5:95)) 119
Table 3.8. Optimised LC-ESI-MS/MS parameters, fraction patterns, molecular weight (Mw), collision energy (CEn) and retention times (RT) of the standard solution of the analytes121
Table 3.9. Identified products of the glycosylated product of the LgtC catalysed reaction when Rgal was used as an acceptor 124
Table 3.10. Enzyme kinetic values of UDP-Gal in various assay conditions (n=3) 130
Table 3.11. The gradient of the optimised HPLC method used in the system suitability and K_m experiments 134
Table 3.12. Optimised HPLC gradient for the separation of UDP-Gal, UDP, Rgal and unknown product form LgtC catalysed reaction..... 137
Table 3.13. Gradients examined with the ZIC-*p*HILIC column at 60 °C..... 139
Table 3.14. Gradients examined with the amide column at 30 °C 140
Table 3.15. Phosphatase coupled biochemical assay set up with LgtC 142
Table 4.1. Structures of examined thiazolidinones..... 147
Table 4.2. Structures of the hit compounds and percentage inhibition at 50 µM against LgtC by HPLC and percentage fluorophore displacement by LDA..... 153
Table 4.3. The functional groups (R_1 and R_2, see the positions from the Scheme 4.1 and Scheme 4.2), yields and purities of the synthesised PP analogues 156
Table 4.4. K_i values (µM) determined at 0, 0.001 and 0.01 % of TX-100 against LgtC and β1,4-GalT by PCA..... 161
Table 4.5. Average particle size of PP analogues (*n*=3) in 13 mM HEPES buffer (pH 7.0) including DMSO (10 % (v/v))..... 163
Table 4.6. Structure of the screened pyrazolones analogues..... 165
Table 4.7. Percentage inhibition by the PCA at 25 µM and K_i values against LgtC and β1,4-GalT in the presence and in the absence of TX-100..... 170
Table 4.8. K_m (µM) and V_{max} (µM/ min) of UDP-Gal in the presence of p5 or p18 at 2.0 µM and 5.0 µM..... 174
Table 4.9. Concentration of the components in the DOS library screening (96 well microplate) 179
Table 4.10. Concentration of the components on the target-focused library screening (384 well microplate) by the LDA 180

Table 4.11. Concentration of the components on the target-focused library screening (96 well microplate) by the PCA 180
Table 4.12. λ_{max}, HPLC gradient utilised and the purity of the synthesised compounds 191
Table 4.13. HPLC gradients utilised in the determination of the purity of the compounds..... 191

Abbreviations

5FTUDP-Gal	5-(5-formyl-2-thiophene)-UDP-Gal, fluorophore
5FTUDP-Glc	5-(5-formyl-2-thiophene)-UDP-Glc
α1,3-GalT	α1.3-galactosyltransferase from *Bos Taurus*
ACN	acetonitrile
A. pleuropneumoniae	*Actinobacillus pleuropneumoniae*
BSA	bovine serum albumin
B. taurus	*Bos taurus*
Arg	arginine
Asp	aspartic acid
br	broad singlet
CAZy	Carbohydrate Active Enzymes Database
C. difficile	*Clostridium difficile*
CE	capillary electrophoresis
CEL	hen egg-white lysozyme
CEn	collision energy
CIP	calf intestinal alkaline phosphatase
C. jejuni	*Campylobacter jejuni*
cLogP	calculated LogP
CMC	critical micelle concentration
CMP-NeuAc	cytidine 5-monophospho-*N*-acetylneuraminic acid
CPD	cysteine protease domain
Cys	cysteine
d	doublet
DAD	diode array detector
DATH	2,4-diacetamido-2,4,6-trideoxyhexose
dd	doublet of doublets
DLS	dynamic light scattering
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DOS	diversity-oriented synthesis
DTT	dithiothreitol
E	enzyme

E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
ELLA	enzyme-linked lectin assay
ER	endoplasmic reticulum
ESI	electrospray ionisation
F	fluorophore
FACE	fluorophore-assisted capillary electrophoresis
FDA	U.S. Food and Drug Administration
FI	fluorescence intensity
FP	fluorescence polarization/anisotropy
FRET	fluorescence resonance energy transfer
Gal	galactose
Gal-β1-4-GlcNAc	*N*-acetyllactosamine
GalT	galactosyltransferase
GCS	glucosylceramide synthase
GDP	guanosine diphosphate
GDP-Fuc	guanosine diphosphate fucose
GDP-Man	guanosine diphosphate mannose
Glc	glucose
GlcNAc	*N*-Acetylglucosamine
GlcT	glucosyltransferase
GT	glycosyltransferase
GTB	blood group B, α-1.3-galactosyltransferase from *Homo sapiens*
GTP	guanosine triphosphate
HCl	hydrochloric acid
HILIC	hydrophilic interaction liquid chromatography
HEP	L-Glycerol-D-manno-heptose
HEPES	*N*-(2-Hydroxyethyl)-piperazine-*N*'-ethanesulfonic acid
H. influenzae	*Haemophilus influenzae*
HPLC	high performance liquid chromatography
H. sapiens	*Homo sapiens*
HTS	high-throughput screening
I	inhibitor
IMAC	immobilised metal affinity chromatography

IEC	ion-exchange chromatography
IPC	ion-pair chromatography
IPTG	isopropylthio-β-galactoside
kDA	kilodalton
KDO	3-deoxy-D-manno-2-octulosonic acid
k_{cat}	enzymatic turnover number
K_i	inhibition constant for competitive inhibitor
K_i'	inhibition constant for mixed inhibitor
K_m	Michaelis-Menten constant
LC	liquid chromatography
LDA	ligand displacement assay
LDH	lactate dehydrokinase
Leu	leucine
LgtA	β1,3-N-acetylglucosaminetransferase
LgtB	β1,4-galactosyltransferase
LgtC	α1,4-galactosyltransferase from *Neisseria meningitidis*
LgtF	β1.4-glucosyltransferase
LNT	lacto-N-neotetraose (Gal-β1,4-GlcNAc-β1,3-Gal-β1,4-Glc)
LOQ	limit of quantification
LogP	compound's hydrophilicity
LOS	lipooligosaccharide
LPS	lipopolysaccharide
m	multiplet
Me	methyl
MeOH	methyl alcohol
MOPS	3-(N-morpholino) propanesulfonic acid
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MurG	GlcNAc transferase
Mw	molecular weight
NAD^+	β-nicotinamide adenine dinucleotide
NADH	β-nicotinamide adenine dinucleotide, reduced form
NaOH	sodium hydroxide
NBD	4-nitrobenzo-2-oxa-1,3-diazole

NDP	nucleoside diphosphate
N. gonorrhoeae	*N. gonorrhoeae*
NGT	N-glycosyltransferase from *Actinobacillus pleuropneumoniae*
N. meningitidis	*Neisseria meningitidis*
NMR	nuclear magnetic resonance spectroscopy
NP-LC	normal phase liquid chromatography
OGT	O-GlcNAc transferase
PCA	phosphate coupled assay
PDI	polydispersity index
PEP	phosphoenolpyruvate
Ph	phenyl
PK	pyruvate kinase
pKa	acid dissociation constant
*p*NPβGal	*p*-nitrophenyl-β-galactoside
ppGalNAcTs	*N*-acetyl-α-galactosaminyltransferases
QED	quantitative estimate of drug-likeness
Rgal	resorufin-β-D-galactopyranoside
RP-LC	reversed phase liquid chromatography
RSD%	relative standard deviation
R	resolution
R^2	regression
RT	retention time
s	singlet
S/B	signal to background ratio
Std	standard deviation
T	tailing factor
TBAHS	tetrabutylammonium hydrogen sulphate
TcdB	bacterial toxin B from *Clostridium difficile*
TRIS	tris (hydroxymethyl)-aminomethane
Trp	tryptophan
TX-100	triton X-100 (polyoxyethylene octyl phenyl ether)
S	substrate
S/B	signal to baseline ratio
UMP	uridine monophosphate

Und-P	undecaprenol phosphate
UDP-2FGal	UDP-2-deoxy-2-fluoro-galactose
UDP	uridine diphosphate
UDP-Gal	uridine diphosphate galactose
UDP-GalNAc	uridine diphosphate *N*- acetylgalactose
UDP-Glc	uridine diphosphate glucose
UDP-GlcA	uridine diphosphate glucuronic acid
UDP-NAcBac	uridine diphosphate-di-*N*-acetylbacillosamine
UDP-GlcNAc	uridine diphosphate *N*-acetylglucosamine
UPD-Xyl	uridine diphosphate xylose
UTP	uridine triphosphate
v	initial velocity
V_{max}	maximum reaction velocity
Z' factor	screening window coefficient
λ_{em}	emission wavelength
λ_{ex}	excitation wavelength
λ_{max}	maximum absorbance wavelength

Acknowledgements

Firstly I wish to thank my supervisors Dr Gerd Wagner and Dr Barry Panaretou for the guidance through the entirety of my PhD and the invaluable support and advice that they have offered.

My thanks also go to King's College London for the studentship, and King's College London and University of East Anglia for financial support enabling me to carry out this PhD.

Special thanks to Dr Andy Evitt for his fantastic guidance in my research work, to Dr James Heaton for his guidance in the analytical work, to Dr Sebastian Gehrke for the synthesis of the pyrazolone library and to Dr Lauren Tedaldi for the synthesis of the UDP-sugar fluorophores and for proofreading this thesis.

I wish to extend my gratitude to the following: Anna Caldwell from the mass spectrometry faculty for assistance with LC-MS/MS experiments, Professor Jayne Lawrence and Dr Laila Kudsiova for assisting in the DLS experiments, Dr Norman Smith for allowing me the use of his HPLC.

I would like to thank the EPSRC National Mass Spectrometry Service Centre (Swansea), for the recording of mass spectra. The plasmids for LgtC and β1,4-GalT were generous gifts from Warren Wakarchuk (Toronto) and Christelle Breton (Grenoble). Additional thanks go to Dr Jon Cuccui and Professor Brendan Wren (London School of Hygiene & Tropical Medicine) for providing NGT, Professor Klaus Aktories and Dr Thomas Jank (Institute of Experimental and Clinical Pharmacology and Toxicology, University of Freiburg, (Germany), for providing TcdB and Dr David Spring (University of Cambridge) for providing the DOS library.

I wish to thank all past and present members of the Wagner group: Andy, Ferdinand, George, Karine, Lauren, Sarah and Sebastian for their support and guidance over the course of my studies.

1. Introduction

Carbohydrates (also called glycans) are the most diverse biological compound class in nature. Initially, carbohydrates were identified as energy storage molecules (e.g. cellulose and glycogen), but since 1960s carbohydrates have also been found to be involved in numerous biological processes such as pathogen infections, inflammation, cell adhesion and cancer.[1] Carbohydrates are represented in many forms from single saccharides to complex polysaccharides and glycoconjugates (when mono- or disaccharide is covalently linked to a non-carbohydrate moiety such as protein or lipid[2]). Glycosyltransferases and glycosidases are two major classes of carbohydrate-active enzymes that modify these carbohydrate structures.[3] Glycosyltransferases catalyse the glycan synthesis (e.g. oligosaccharides) whereas glycosidases hydrolyse glycosidic linkages.

Specific inhibitors of glycosyltransferases could serve as therapeutic leads in drug discovery and could be utilised as useful tools in analysing glycan function and assembly. Therefore, robust and sensitive assays are essential for identifying novel inhibitors, monitoring enzymatic activity and inhibitor potency. The main focus of the thesis is to develop both a high-throughput assay for inhibitor identification and a biochemical assay for investigation of activities of GTs.

This chapter will introduce glycosyltransferases, their classification, mechanisms, biological function and potential drug targets. This will be followed with a brief review of the analytical techniques utilised for GT characterisation. Glycosidases, the second main carbohydrate-active enzyme class, are also briefly introduced.

1.1. Glycosyltransferases

Glycosyltransferases (GTs) represent a large family of enzymes that are involved in the biosynthesis of glycoconjugates, oligosaccharide and polysaccharides.[1, 4, 5] These enzymes catalyse the transfer of a sugar moiety from an activated donor sugar substrate onto specific acceptor molecules forming glycosidic bonds (Figure 1.1).[5] The acceptors are most commonly saccharides but can also be lipids, natural products, DNA, proteins or glycoconjugates.[5] Generally, GTs elongate the glycan chains by transferring sugar moieties producing a preferred acceptor substrate for another GT. The glycosidic linkage between each sugar moiety can have two anomeric configurations (α or β),[5] therefore, the potential number of structures that can exist is vast.

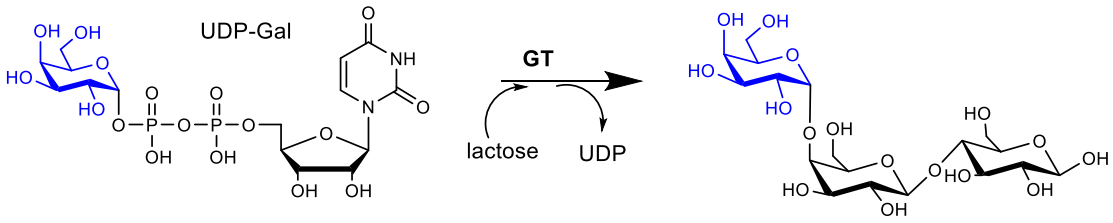
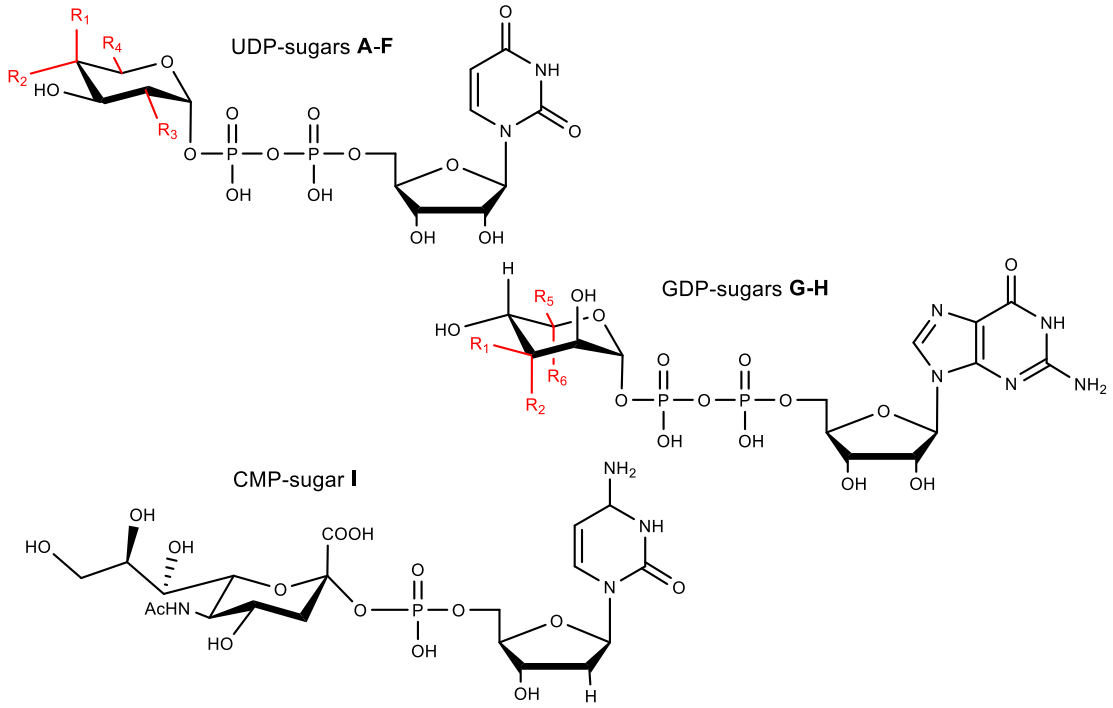


Figure 1.1. Glycosyltransferase catalysed reaction

Most of the GTs are nucleotide sugar-dependent enzymes and they are usually referred to as Leloir enzymes.[6, 7] The donor substrates consist of a sugar moiety which is linked to a nucleotide *via* a mono-or diphosphate. The high-energy bond between the anomeric centre of the sugar ring and phosphate activates the sugar for the transfer reaction. In mammalian metabolism there are nine main nucleotide donors (Table 1.1): uridine diphosphate galactose (UDP-Gal), uridine diphosphate glucose (UDP-Glc), uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc), uridine diphosphate *N*- acetylgalactose (UDP-GalNAc), uridine diphosphate xylose (UDP-Xyl), uridine diphosphate glucuronic acid (UDP-GlcA), guanosine diphosphate mannose (GDP-Man), guanosine diphosphate fucose (GDP-Fuc) and cytidine 5'-monophospho-*N*-acetylneuraminic acid (CMP-NeuAc).[3] The number of nucleotide donors is higher in plants and bacteria.[8]

Table 1.1. Nucleotide sugar donors in mammalian metabolism



Nucleotide sugar	R1	R2	R3	R4	R5	R6
UDP-Gal (A)	-OH	-H	-OH	$-CH_2OH$		
UDP-Glc (B)	-H	-OH	-OH	$-CH_2OH$		
UDP-GlcNAc (C)	-H	-OH	-NHAc	$-CH_2OH$		
UDP-GalNAc (D)	-OH	-H	-NHAc	$-CH_2OH$		
UDP-Xyl (E)	-H	-OH	-OH	-H		
UDP-GlcA (F)	-H	-OH	-OH	-COOH		
GDP-Man (G)	-OH	-H			$-CH_2OH$	-H
GDP-Fuc (H)	-H	-OH			-H	$-CH_3$
CMP-NeuAc (I)						

GTs are highly specific for both the donor and the acceptor substrates, and generally, do not tolerate structural modifications of their natural substrates. However, some GTs can accept unnatural substrates and the reactions can occur in a non-stereospecific manner.[9-13] Recently, substrate engineering studies have shown that a bifunctional α2,3/α2,8 sialyl transferase CstII from *Campylobacter jejuni* (*C. jejuni*) is able to use unnatural donors and acceptors. CstII strain OH4384 naturally catalyses the transfer of sialic acid moiety from CMP-NeuAc to the terminal position 3' hydroxyl of terminal lactose-containing acceptors. The second step, sialic acid is transferred to the 8' hydroxyl of the α2,3'-sialylated lactose.[14] Lairson and co-workers investigated donor flexibility of CstII with α-linked *para*-nitrophenyl (pNP) sialic acid where it could serve as an alternative donor.[13] They demonstrated that CstII catalysed the transfer of sialic acid with α2,3-linkage but only in the presence of the natural donor CMP-NeuAc.[13]

Recent investigations with a derivative of galactose, (β-linked pNP-Gal), demonstrated that it served as a specific acceptor for CstII, and also for bovine α1-3-galactosyltransferase (α1,3-GalT)[12] which naturally catalyses the transfer of galactose from UDP-Gal to terminal *N*-acetyllactosamine-containing glycoconjugates.[15] A non-stereospecific outcome of the GT catalysed reaction has also been reported. Acceptor specificity studies of LgtC (α1,4-galactosyltrasferase) demonstrated that the transfer of galactose to unnatural acceptors formed α1,2/3/4 linkages.[12] LgtC naturally forms an α1,4-linkage between donor sugar and acceptor molecule.[16] These properties can be exploited as tools for the enzymatic synthesis of various oligosaccharides.

1.1.1. Classification of glycosyltransferases

Glycosyltransferases have been classified into families based on their amino acid sequence similarities along with the stereochemical outcome of their reactions and three-dimensional structure.[5, 17] To date 94 GT families based on their amino acid similarities are listed in the Carbohydrate Active Enzymes Database (CAZy).[5, 18] The majority of the known GTs can be divided into two classes with the GT-A and GT-B folds based on their three-dimensional crystal structures.[5, 18] However, no correlation is found between the GT folds and the stereochemical outcome of the catalysed reaction. GT-A and GT-B folds are formed of two Rossmann-like folds which is typical for nucleotide-binding proteins.[1, 5] The Rossmann-like fold consists of βαβ units, usually 6 parallel β- strands linked to 4 α- helices.[19]

The GT-A proteins contain two closely associated Rossmann-like folds, as a single domain.[5] The first GT-A fold was identified when the crystal structure of SpsA from *Bacillus subtilis* was described in 1999 (Figure 1.2, left).[20] Most GT-A folded GTs have an Asp-x-Asp (or DXD, x is any amino acid) motif in the catalytic site, which is involved in metal ion binding, such as Mn^{2+} or Mg^{2+}, and is essential for enzymatic activity.[5] The GT-A fold also contains a small flexible loop region that changes conformation during the catalytic cycle by creating an acceptor binding site.[21] The first described GT crystal structure was the β-glycosyltransferase from bacteriophage T4 in 1994.[22]

The GT-B proteins contain two Rossmann-like folds, however, in this case the two domains are less associated and the active site is predicted as lying in the cleft (Figure 1.2, right).[5] Instead of using a divalent metal ion to stabilise the nucleotide phosphate, the joined helices of the subdomain of the Rossmann folds create a binding site for the acceptor.

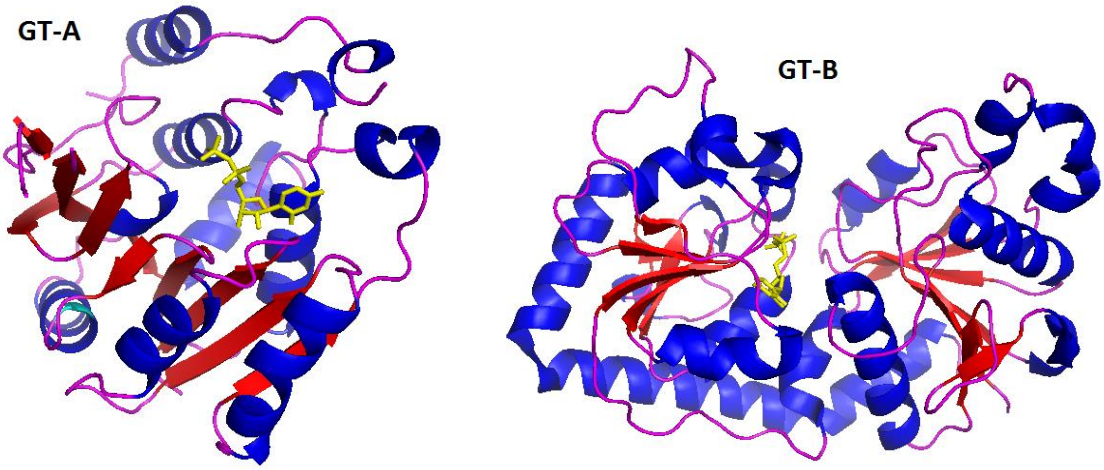


Figure 1.2. Overall folds observed for GTs. The GT-A fold is represented by the inverting GT SpsA from *Bacillus subtilis* in the presence of UDP (Protein Data Bank, PDB: 1QGQ)[20] and the GT-B fold, by bacteriophage T4 β-glucosyltransferase in the presence of UDP (PDB: 1JG7)[23]. α-helixes are in blue, β-strands in red and donor substrates are presented in stick form (yellow)

A GT-C fold has been recently proposed for hydrophobic integral membrane GTs and this fold is proposed to bind lipid phosphate-activated donor sugar substrates.[5] The membrane enzymes are difficult to express and crystallise, and only two crystal structures have been reported.[24, 25] Most GTs with known structure possesses the GT-A, GT-B or GT-C fold, however, some GTs still remain unclassified.

1.1.2. Mechanism of GTs

GTs catalyse the formation of a glycosidic bond between the sugar from the nucleotide donor and the acceptor. The released glycosylated product can have the same or opposite configuration at the anomeric carbon as the donor substrate (Figure 1.3).[5] Functionally, GTs can be separated into retaining or inverting enzymes depending on the outcome of the reaction. The mechanism for inverting GTs is well established while the retaining mechanism is still indistinct.

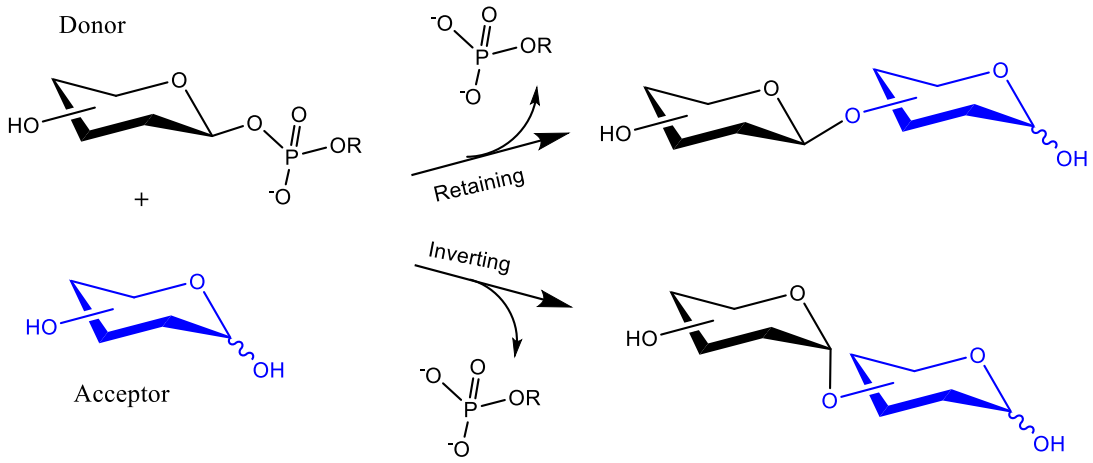


Figure 1.3. Retaining and inverting transfer reactions catalysed by glycosyltransferases[5]

The inverting GTs follow a single S_N2-like displacement mechanism.[5] A base (usually asparagine or glutamic acid) in the active site of the enzyme, protonates the hydroxyl group of the acceptor, assisting its direct nucleophilic attack onto the anomeric carbon of the sugar donor (Figure 1.4, a). Retaining GTs are proposed to use a double displacement mechanism which involves two sequential S_N2-like reactions that lead to a covalently bound glycosyl-enzyme intermediate (Figure 1.4, b).[1, 5] The mechanism requires a nucleophile and divalent cation or charged side chain in an appropriate location within the active site. The first step occurs similarly to the inverting mechanism and the hydroxyl group of the incoming acceptor is probably activated for the nucleophilic attack by the base catalyst of the leaving group or the enzyme.

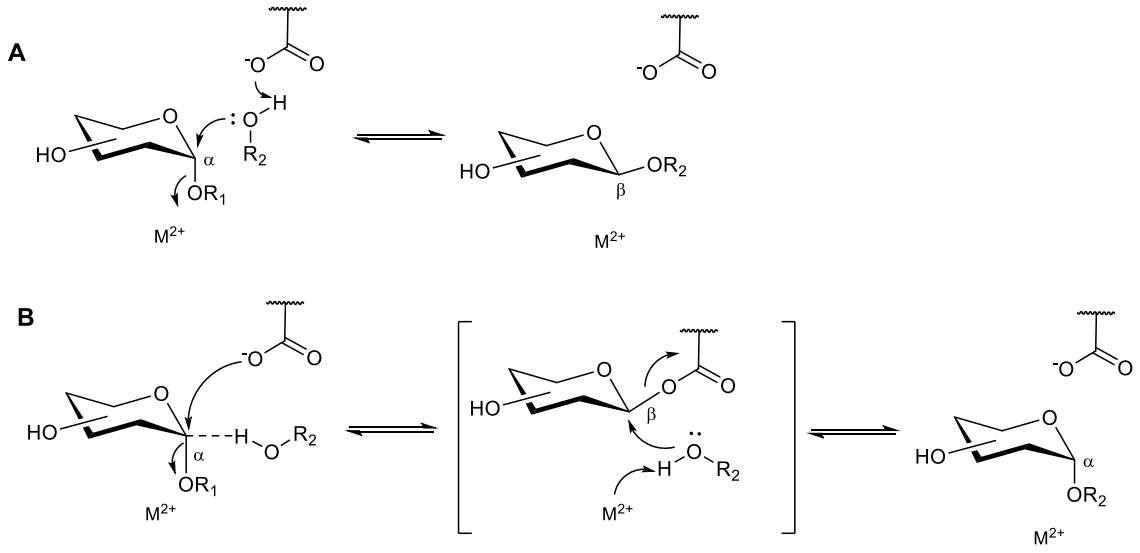


Figure 1.4. Direct-displacement mechanisms of inverting GTs (A) and a proposed double-displacement mechanism for retaining GTs (B). R_1 would generally be a nucleoside mono- or diphosphate of the donor substrate and R_2 would correspond to the acceptor substrate

Alternatively, S_Ni-like and S_Ni mechanisms have been proposed for retaining GTs which involves an ion pair intermediate or oxocarbenium ion-like transition state, respectively.[5, 26] Studies of the reaction mechanism of retaining GTs can be controversial. For example, S_Ni-like and S_Ni mechanisms have been proposed for α1,4-galactosyltransferase (LgtC) based on experimental and computational studies.[16, 27, 28]

1.1.3. Glycosyltransferases in biological systems

The products of GT catalysed reactions such as glycolipids and glycoproteins mediate many essential biological processes. These complex glycan chains are generally located on the cell wall of eukaryotic cells where they play an important role as receptors in recognition events such as cell signalling and cellular adhesion, and provide receptors for viruses, bacteria or hormones.[29] GTs are involved in the biosynthesis of *N*-/*O*-liked glycoproteins and glycolipids in both prokaryotic and eukaryotic organisms.[29-31] Glycosylation in eukaryotic cells is initiated in the endoplasmic reticulum and further modified in the Golgi apparatus.[29] Whereas in prokaryotic cells glycosylation occurs in the cytoplasm and in the periplasm.[32]

Glycosylation plays a significant role for human health and disease.[33-35] Changes in glycosylation can cause under- and over expression of naturally occurring glycans. In diseases such as cancer, glycosylation of the cell surface proteins change from normal and the immune system initiates the construction of abnormal cells when the wrong type of glycans are recognised. For example, *N*-acetylglucosaminyltransferase V (GlcNAc-TV, also known as MGAT5: mannoside-acetylglucosaminyltransferase 5) has been found to be associated in the branching of *N*-linked oligosaccharides.[33, 36] These extensively β1,6-GlcNAc branched glycans create additional sites for terminal sialic acid residues on the cell surface, which can lead to an elevated activity of sialyltransferases in cancerous cells. Many polysaccharide structures, such as sialyl Lewis isomers that are present on *O*- and *N*-linked oligosaccharides, have been found to be associated in various human malignancies. Increasing activity of fucosyltransferases, such as FucT I-VIII which transfers fucose in α 1,2/3/4 or 6-linkages to terminal galactose or GlcNAc residues can lead to an over expression of various terminal glycans, such as sialyl Lewis X and Lewis Y (Figure 1.5).[36] For example, increasing activity of α1,2-fucosyltransferase in human colorectal cancer cells have shown to be linked to the formation of Lewis B (Fuc-α1,2-Gal-β1,

3-[Fuc-α1,4]-GlcNAc-β1,3-Gal-β1,4-Glc) and Lewis Y[37] and, increased activity of β1,6-*N*-acetylglucosaminyltransferase (C2GnT) in breast carcinoma, sialyl Lewis A (NeuAc-α2,3-Gal-β1, 3-[Fuc-α1,4]-GlcNAc) and sialyl Lewis X are present but are not in normal breast epithelial cells.[38]

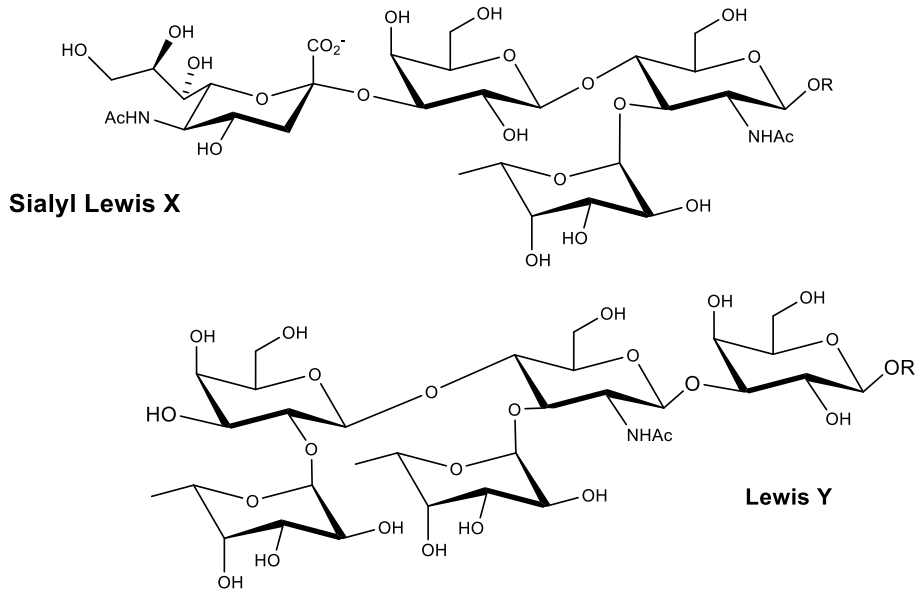


Figure 1.5. Structure of two cancer-associated glycans: sialyl Lewis X and Lewis Y[33]

Almost all bacteria cell walls consist of a peptidoglycan layer which is formed of linear glycan strands cross-linked by short peptides.[39] The peptidoglycan layer is essential for shape and protects the bacteria from turgidity. The outer cell membrane of the Gram-negative bacteria is covered with polysaccharides (Figure 1.6) and this sugar layer is responsible for the protection of the bacteria from the host's immune system, as well as acting as a virulence factor for the infection of the host.[32, 40] Unusual monosaccharides such as di-*N*-acetylbacillosamine ($BacAc_2$), 2,4-diacetamido-2,4,6-trideoxyhexose (DATH), *N*-acetylfucosamine and pseudaminic acid have been found in bacterial glycosylation that are not present in the eukaryotic systems.[30] Both *O*-linked and *N*-linked protein glycosylation pathways are found in bacteria and the first bacterial *N*-glycosylation pathway was described with *C. jejuni*.[41] In *C. jejuni*, five different glycosyltransferases (*i.e.* PglC, PglA, PhlJ, PglH and PglI) are involved in the biosynthesis of heptasaccharide (GalNAc-α1,4-GalNAc-α1,4-(Glc-β1,3)-GalNAc-α1,4-GalNAc-α1,4-GalNAc-α1,3-$BacAc_2$) on a lipid carrier which is then transferred to an asparagine residue of the protein by PglB. In addition, unusual monosaccharides are found in *O*-glycosylation in *Neisseria gonorrhoeae* where DATH containing trisaccharide (Gal-β1,4-Gal-α1,3-DATH) is attached to serine residue.[30]

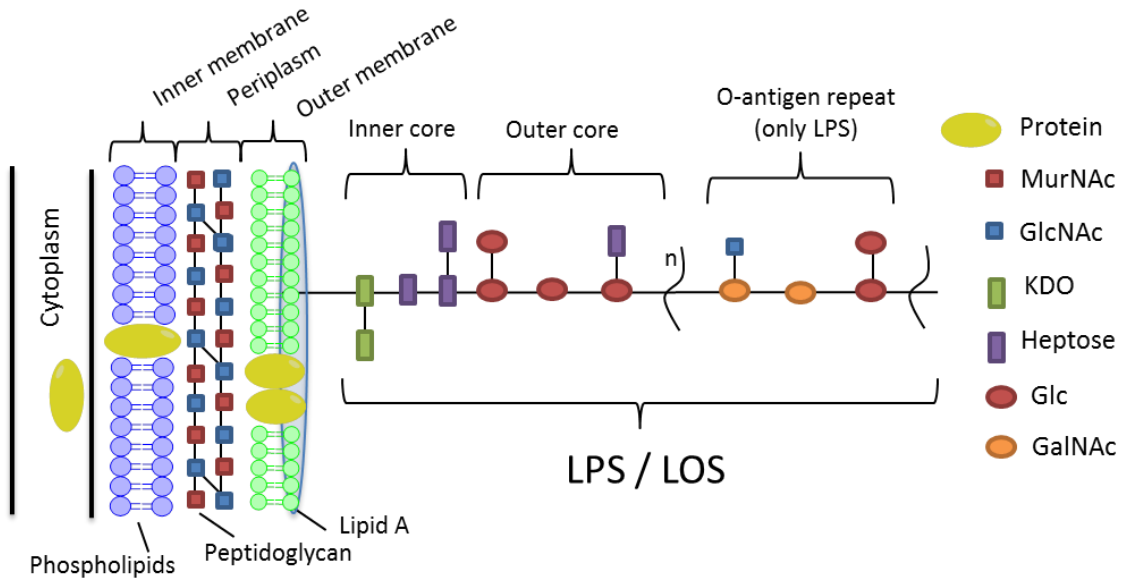


Figure 1.6. Schematic presentation of the cell wall of Gram-negative bacteria (modified from reference[42])

A range of GTs play an important role in the biosynthesis of Gram-negative bacterial lipopolysaccharides (LPSs) and lipooligosaccharides (LOSs). LOS differs from LPS by lacking the *O*-antigen (*O*-polysaccahride) repeating units. LPS and LOS are located in the outer membrane of the bacterial cell surface (Figure 1.6) where they are attached to lipid A which is a key factor for virulence.[40] The outer oligosaccharide core of LPS and LOS is an important factor for the interaction between the bacteria and the environment. For example, various α-chains can be formed in the biosynthesis of LOS in *Neisseria meningitidis* (*N. meningitidis*) that are crucial for bacterial survival in the host. The biosynthesis of meningococcal LOS is initiated in the cytoplasm by the synthesis of an inner oligosaccharide core that is attached to lipid A, containing two KDO (3-deoxy-D-manno-2-octulosonic acid) and two HEP (L-glycero-D-manno-heptose) residues.[43] Synthesis of the outer oligosaccharide core (α-chain) is initiated by LgtF (β1,4-glucosyltransferase) that first adds glucose to the HEP1 residue, followed by the sequential addition of additional sugar residues (Figure 1.7). The completed α-chain is translocated to the periplasmic surface of the inner membrane, transported across periplasm and integrated into the outer membrane.[43] Various GTs are responsible for the sequential addition of sugars to the α-chain. Depending on the phase variable expression of lgtA and lgtC (LgtA and LgtC encoding genes, respectively) different α-chains can be formed.[44, 45] If LgtA (1,3-*N*-acetylglucosaminetransferase) is not expressed then the α-chain can terminate at the Gal-β1,4-Glc-β1,4-HEP1. Alternatively, if LgtC is expressed then addition of α1,4-Gal at the Gal of the Gal-β1,4-Glc-β1,4-HEP1 of LOS occurs and a short α-chain (Gal-α1,4-Gal-β1,4-Glc,

globotriaose) is formed.[46, 47] Synthesis of lacto-*N*-neotetraose (Gal-β1,4-GlcNAc-β1,3-Gal-β1,4-Glc, LNT) occurs in the presence of LgtA, transferring GlcNAc to the terminal galactose followed by LgtB (β1,4 galactosyltransferase) to terminate the LNT structure.[48, 49]

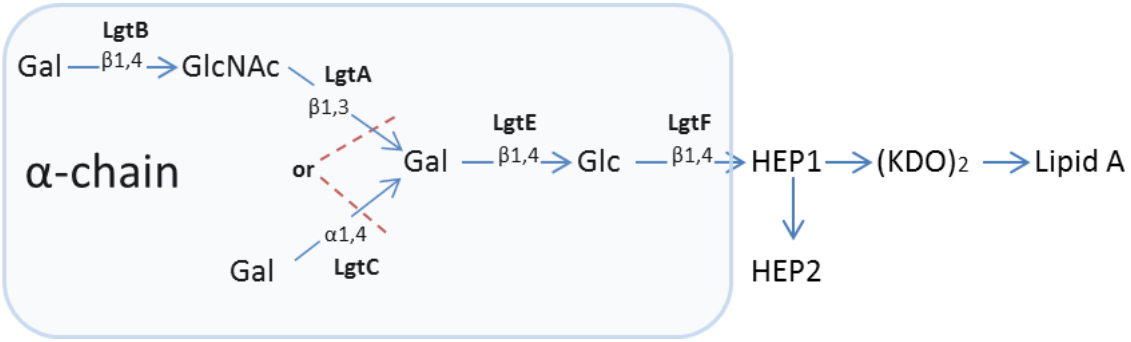


Figure 1.7. LOS of *N. meningitidis*. LgtC transfers galactose to the terminal end of the α-chain if LgtA is inactive (the figure is modified from reference[43])

Globotriaose and LNT chains on *N. meningitidis* LOS are major virulence determinants. These α-chains can be sialylated and they form a capsular polysialic acid structure that camouflages *N. meningitidis* from the human immune system.[46, 47] In addition, sialylation increases survival of *N. meningitidis* in the blood stream that can lead to fatal septicaemia.[50]

1.1.4. Inhibitors of glycosyltransferases

GTs are involved in the biosynthesis of complex glycoconjugates that participate in fundamental properties of cells such as signalling, immune response, microbial adhesion and infection. Because the glycosylation is essential for the survival of organisms, GT inhibitors are potential chemical tools for studying the behaviour of these enzymes, such as their mechanism, their role in glycan synthesis and also as potential lead compounds in drug discovery. Different strategies have been used in order to design potent GT inhibitors. Generally, the designed structure analogues are based on donor or acceptor substrates and potency as an inhibitor against a certain GT is investigated by a biochemical assay.[51] In addition, three-dimensional structures of the enzymes are used to design potential inhibitors based on docking simulations.[52-54] Another approach for inhibitor discovery is high-throughput screening (HTS) of large compound libraries. HTS has been shown to be an effective technique for the discovery of small molecule inhibitors and novel GT inhibitors have been identified (Figure 1.8, D-F).[55-57]

Different types of GT inhibitors have been developed to date, including acceptor and donor analogues, transition-state analogues and small molecule inhibitors.[51, 58, 59] A disadvantage of the donor analogues (e.g. carbasugar analogues of nucleoside diphosphate sugars, Figure 1.8, A).[60]) is a poor cell membrane permeability due to their polar and charged nature, therefore, they generally cannot be used for *in vivo* studies. Recently, membrane permeable inhibitors of GTs have been developed.[61, 62] A sugar analogue of GlcNAc, 2-naphthyl 2-butanamido-2-deoxy-1-β-D-glucopyranoside (Figure 1.8, B) was found to be a specific inhibitor of β1,4-Galactosyltransferase in cell homogenates (human airway cells; H292, human colonic cancer cells; CaCo-2 and mouse lymphocytic cells; MOPC).[61] Brown and co-workers discovered an acceptor analogue of GlcNAc-β1,2-Man-β1,6-Man. Deacetylated GlcNAc-β1,3-Gal-β-*O*-naphthalenemethanol inhibited the activity of β1,4-Galactosyltransferase T1 and the *in vivo* studies indicated the formation of sialyl Lewis x was blocked in U937 monolytic leukemia cells (Figure 1.8, C).[62]

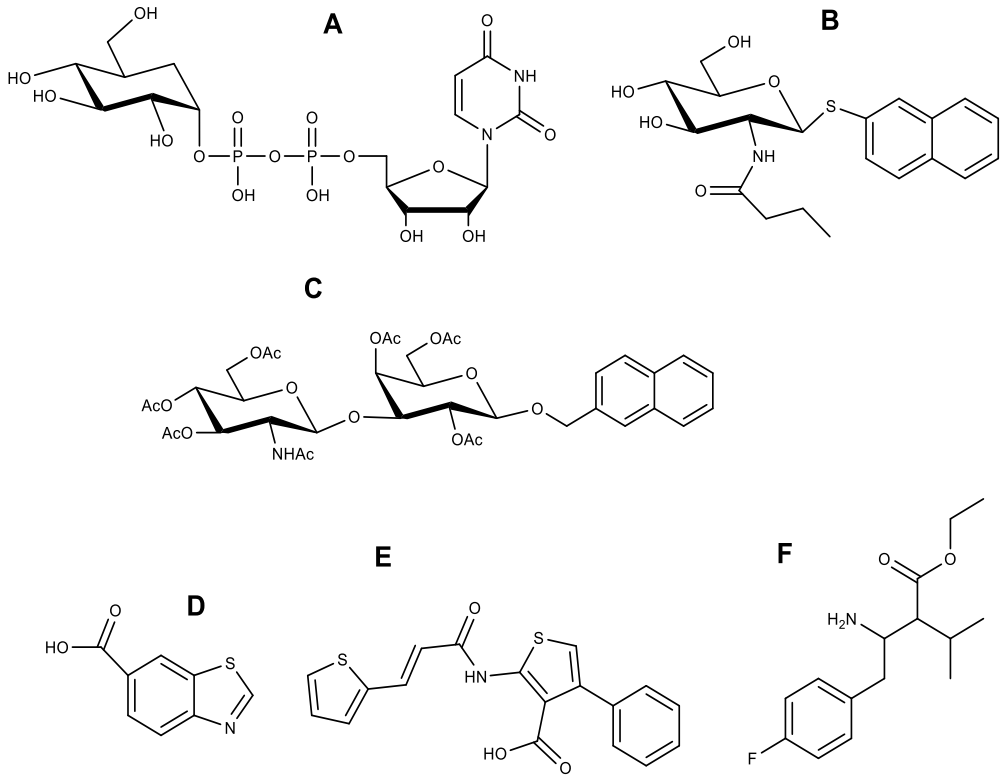


Figure 1.8. A) UDP-carba-Glc, a donor substrate analogue B) An acceptor analogue, 2-naphthyl 2-butanamido-2-deoxy-1-β-D-glucopyranoside inhibitor of β1,4-GalT[61], C) An acceptor analogue, per-O-acetylated GlcNAc-β1,3-Gal-β-*O*-naphtalenemethanol, inhibitor of β1.4-GalT T1[62], D-F) small molecular inhibitors of *O*-GlcNAc transferase identified by HTS[56]

A small number of drug-like GT inhibitors have been discovered and some of these inhibitors have reached clinical study stage.[59, 63] Currently in clinical use, Miglustat, *N*-butyl-deoxynojirimycin (Figure 1.9), has been used for treatment of type 1 Gaucher's disease which is a glycosphingolipid lysosomal storage disorder.[64] Gaucher's disease is caused by a lack of glucocerebrosidase which leads to an accumulation of glucosylceramide in several organs. Miglustat is an imino sugar inhibitor that lowers the activity of glucosylceramide synthase which catalyse the first step in the biosynthesis of glucosphingolipids.[64, 65]

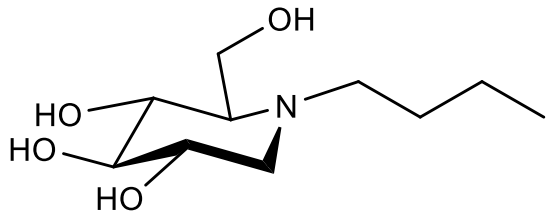


Figure 1.9. Chemical structure of Miglustat[65]

Searching for potent inhibitors of GTs is challenging due the low binding affinity of glycosyl donors and acceptors to GTs, the lack of three-dimensional protein structures and the complex reaction mechanism, which includes a transition state where a nucleotide sugar, an acceptor and often, a metal ion are involved. In addition, the lack of sensitive, robust and operationally simple compound screening methods due to non-chromophore containing substrates, further complicate this process.

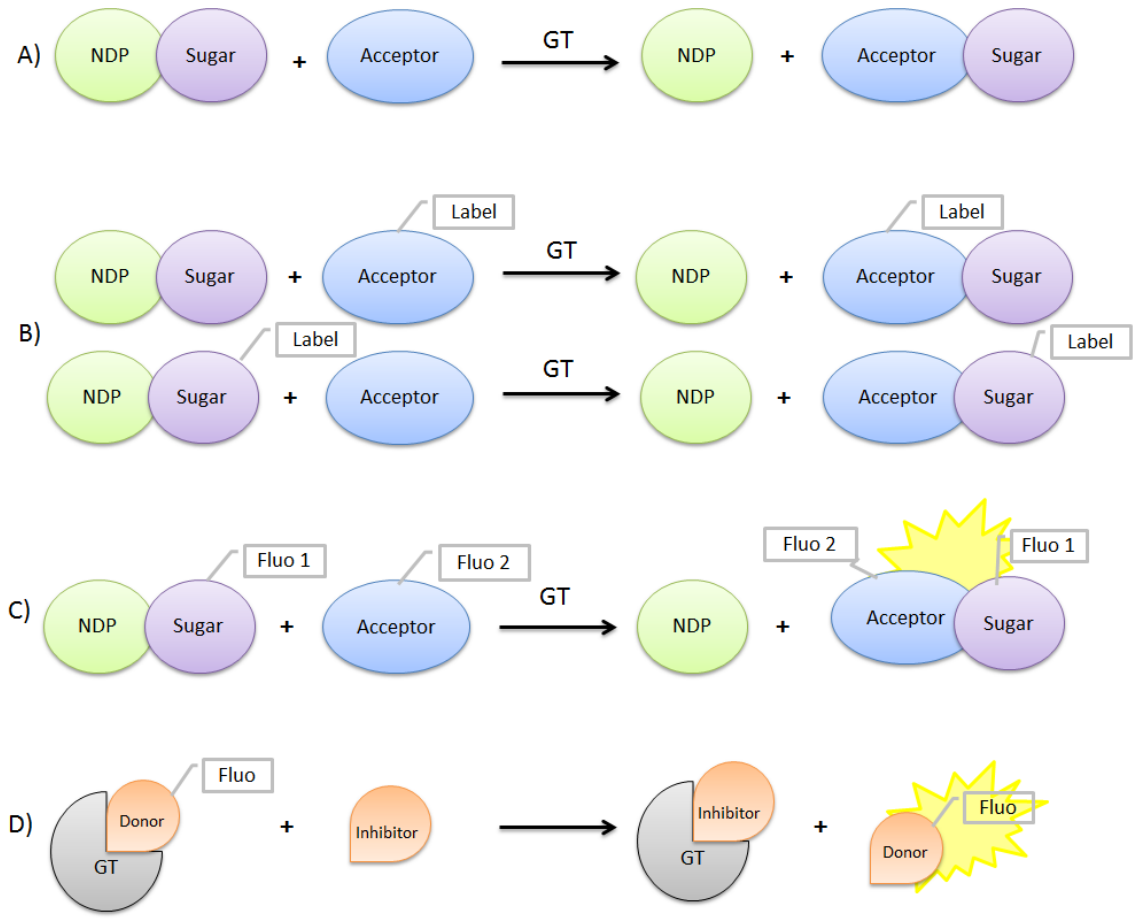
1.2. Glycosyltransferase assays

Numerous diverse assays have been developed to study GTs.[63, 66] These assays allow the determination of GT activity during and after enzyme purification in cellular systems, and can be used for mechanistic studies, for high through-put screening of GT inhibitors as well as for the directed evolution of GTs.

GT assays can be designed either as biochemical or binding affinity assays.[63] The biochemical assays are carried out with saturated substrate concentrations and progress of the enzymatic reaction can be monitored either by the depletion of the substrates (i.e. sugar-nucleotide or acceptor) or the formation of the products (i.e. nucleoside diphosphate (NDP) or glycosylated acceptor). These assays provide essential information about the biological activity of the used

GT. Therefore, the biochemical assays are generally used for inhibitor validation and substrate specificity studies. A disadvantage of the biochemical assays is the poor availability of many substrates, particularly acceptors. In contrast to the biochemical assay, the binding assays (Scheme 1.1. D)[63] are designed to monitor the binding affinity of a labelled donor substrate. The acceptors are not required in this assay format, and therefore, the conversion of the donor to product is not monitored. The binding assay can be developed as HTS assays where the binding affinity of small molecular inhibitors can be monitored effectively without acceptors. The advantage of the binding affinity assays compared to the biochemical assays is the applicability for HTS of inhibitor candidates, cost effectiveness and simple assay protocols.

Natural GT acceptor substrates are generally (poly) saccharides which do not contain a chromophore or a fluorophore. However, the natural GT catalysed reaction can be monitored by detecting the formation of the glycosylated product (primary product) by mass spectrometry (MS) or the NDP (secondary product) with various detection methods, e.g. coupled enzymatic assays, chromatographic methods with UV detection, chemosensors and immune assays (Scheme 1.1. A). Methods such as high performance liquid chromatography (HPLC), capillary electrophoresis (CE) or ion exchange chromatography with a spectrophotometric detection are generally used for the monitoring of the depletion of the sugar donor or formation of the NDP, due to the UV-active nucleobase. If the method is based on detection of the NDP then it is essential to determine the enzymatic or chemical hydrolysis of the sugar donor under the assay conditions. Another approach for monitoring GT activities is to use labelled donors or acceptors which allow the detection of the primary product (Scheme 1.1. B). The labelled product can then be detected with spectrophotometric techniques (UV or fluorescence) or radiochemical methods (scintillation counting). GT reactions can also be monitored by FRET (fluorescence resonance energy transfer), a technique where both the donor and the acceptor are fluorescently labelled (Scheme 1.1. C). A ligand displacement assay is based on a fluorescent donor analogue that binds to the active site which can be then displaced with a high affinity inhibitor (Scheme 1.1. D). The ligand displacement assay is particularly designed for inhibitor screening.



Scheme 1.1. Principles of assays used for determination of GT activity and ligand binding. A) Native GT reaction; NDP can be detected by HPLC/UV, coupled enzymatic assay, immunodetection or chemosensor and primary product can be detected by MS, B) Labelled donor or acceptor: the product can be detected by spectrophotometric or radiochemical methods, C) FRET based assay principle, D) fluorescence-based ligand displacement assay for detection GT inhibitors (modified from reference[63])

Monitoring the GT reactions is challenging due to the complexity of the reaction which occurs in the presence of multiple, colourless substrates in low concentrations. Therefore, GT assays require a sensitive detection method for the monitoring of small concentration changes in the GT catalysed reactions. Limitations for the development of GT biochemical assays are the poor availability of natural substrates as well as labelled substrates and GTs may not be able to turn over unnatural substrates with bulky fluorescent substituents with useful reaction rates. Ideally, GT assay that monitor the reaction progress in real time are particularly valuable. In addition, accuracy, reproducibility and operational simplicity are important factors for an effective GT biochemical assay. Due to the complexity of the GT reaction a range of different assays have been developed to achieve these requirements. Therefore, GT assays will be introduced and their general advantages and disadvantages, and applicability to HTS are evaluated.

1.2.1. Non fluorescence-based assays

Radiometric assays are the most sensitive techniques for monitoring GT activities.[66] Typically the radiometric assays involve incubation of the enzyme with radiolabelled donor substrate and acceptor. After the desired incubation time the reaction is stopped and unreacted radiolabelled substrates are then removed from the radiolabelled glycosylation product. This step usually requires chromatographic techniques (*i.e.* ion-exchange chromatography, electrophoresis[66]) and the radiolabelled glycosylated product is quantified by scintillation counting.[67] Due to the separation and the washing steps, radiometric assays are difficult to adapt for HTS. Recently, von Ahsen and co-workers develop a miniaturised scintillation proximity assay with fucosyltransferase VII.[68] The acceptor molecules were coated with scintillat-impregnanted microspheres which emit light when stimulated with radiolabelled molecules. Therefore, the reaction was monitored by incubating the coated acceptors with radiolabelled donor substrates. The assay was then applied to HTS and nearly 800000 compounds were screened against human fucosyltransferase VII.[68] Due to the detection technique, washing steps were not required. However, a significant disadvantage of this assay is the radioactive waste and the requirement for radiolabelled substrates.

Many chromatographic methods have been used to determine activities of GTs and the most common techniques are HPLC[69-72] and CE.[73-76] A GT-catalysed reaction can be monitored without additional modifications by detecting the UV-active nucleobase containing sugar donor or/and the nucleobase containing secondary product. In this approach, when the detection is not based on monitoring the primary reaction product, it is imperative to determine the amount of hydrolysis of sugar nucleotide donor. Alternatively, the acceptors can be post- or pre- derivatised with UV-active or fluorescent moieties which then allow the quantitative analysis of the primary product. Hayashi and co-workers utilised commercially available fluorescent acceptors C6-4-nitrobenzo-2-oxa-1,3-diazole (NBD)-ceramide and C6-NBD-glucosylceramide to monitor the activity of GlcT and GalT respectively.[71] The fluorescent products were then separated from the donor substrates and quantified with a HPLC-based method. Monegal and co-workers developed a fluorophore-assisted capillary electrophoresis method to monitor α1.3-GalT activity.[74] The glycosylated product was post-derivatised with 8-aminonaphthalene-1,3,6-trisulfonic acid that allowed both UV-detection and fluorescence detection of the product. In this assay format the GT catalysed reaction occurs in the presence of native substrates which is an advantage compared to the assay where pre-derivatised

acceptors, which may hamper the reaction progress, are used. The chromatographic methods, particularly CE with fluorescence detection,[75] are very sensitive. However, the major disadvantage is the time-consuming procedures because the assays are performed in a discontinuous manner.

Mass spectrometry (MS) offers an alternative as detection methodology in GT assays because it is applicable to monitoring the turnover of naturally occurring GT substrates. MS detection has been used for inhibitor screening[77] and characterisation,[78, 79] and also in HTS format. The HTS format has been employed for studying potential substrates for bovine β1,4-GalT and a plant glucosyltransferase, UGT72B1 form *Arabidopsis thaliana.*[80] This method provided a rapid and reliable determination of kinetic parameters for these two enzymes and the assay format could be suitable for other GTs.

An equivalent proton is released during the transfer of sugar molecule from sugar-nucleotide donor to the acceptor, causing a change in pH which can be detected by using pH indicators. Deng and co-workers developed an assay by exploiting this approach by detecting the pH change with phenol red pH indicator.[81] The pH assay was applied to study the substrate specificities of three different galactosyltransferases. Palcic and Persson also developed a similar pH based assay with pH indicator bromothymol blue.[82] They employed the assay in HTS format and screened 350 mutants of the blood group α1.3-galactosyltransferase (GTB). Two novel mutants were identified and the characterisation investigations with radiochemical assay displayed increased activity compared to the wild-type GTB. Advantages of these two assays are that labelled substrates are not required and the applicability to HTS and other GTs.

Immunological methods are also applied to HTS. Hang and co-workers developed an enzyme-linked lectin assay (ELLA) for polypeptide *N*-acetyl-α-galactosaminyltransferases (ppGalNAcTs) and they screened a small urine-based library against ppGalNAcTs.[83] A disadvantage of such immunological methods is the availability of specific antibodies for each reaction product.

Coupled enzyme assays are performed with natural GT substrates and generally on a microplate, therefore, they are applicable techniques for HTS. Gosselin and co-workers assayed GT activities by coupling the secondary reaction product, UDP, *via* pyruvate kinase and lactate dehydrogenase to NADH oxidation.[84] The reaction progress was proportional to the decrease in absorbance at 340 nm. This continual coupled assay was developed for the

investigation of four different galactosyltransferases. The amount of UDP (or UMP) has also been detected by a coupled reaction with phosphatase which releases inorganic phosphate.[85] The inorganic phosphate can be then detected with colorimetric reagents. Advantage of these coupled assays is that they can be applied to a GT reaction that produces UDP or UMP. Conversely, these types of assays may consume large quantities of valuable reagents when they are applied to large compound library screening. In addition, both reported coupled assays utilise UV-detection which may not be enough for detecting small concentration changes in the GT-catalysed reactions.

1.2.2. Fluorescence-based assays

Fluorescence is a phenomenon where molecules absorb a photon and after a short time they emit a photon which has a longer wavelength and lower energy.[86] Fluorescence intensity (FI), fluorescence anisotropy/polarisation (FP), fluorescence resonance energy transfer (FRET) can be used to monitor glycosylation when suitable fluorescently labelled substrates are found.[63] FI is directly related to changes in fluorophore concentration and is typically used in enzyme assays to monitor substrate turnover.[86] FP provides information on molecular orientation and is typically used for ligand binding reactions. The degree of polarisation is determined from measurements of fluorescence intensities parallel and perpendicular with respect to the plane of linearly polarised excitation light.[87] FRET is a distance-dependent interaction between the electronically excited states of two fluorophores in which excitation is transferred from a donor molecule to an acceptor molecule.[87] FRET can be used to assay interactions of biomolecules over a 1-5 nm scale.[86, 88] Fluorescence quenching can occur by a variety of different processes. For example, molecules in solution, upon contact with the fluorophore can cause a quenching by deactivating the excited photon and emitted light is released as heat instead, or the fluorophore can form a complex with a quencher which has a different absorption spectrum.[87] In addition, high fluorophore concentrations or impurities in sample can cause inner-filter interference.[88] The inner filter is caused by a very high absorbance which prevents the excitation light penetrating deeply enough to the point at which the detection optics are focused.

Fluorescence methods are generally performed on microplates and, because the read-outs are volume-independent, the assays can be developed in low reaction volumes and miniaturised. Additionally, fluorescence detection is more sensitive than UV-detection and does not require expensive instruments. However, an identical emission wavelength of the labelled substrates and tested inhibitors or possible impurities may cause false results in assays that are developed for inhibitor screening. This interference can be prevented by developing a long emission wavelength labelled substrates because most of the compounds in compound collections generally are fluorescent in the blue-green range.

Fluorescence-based ligand displacement assays that monitor ligand binding are an attractive technique for large compound library screenings due to low reaction volume and sensitivity. Fluorescent donor analogues are essential to obtain detectable signals as naturally occurring sugar nucleotides do not contain a fluorophore. However, the assay protocols are generally very simple because acceptors are not required. Walker's group demonstrated the effectiveness of a ligand displacement assay with two GlcNAc transferases, MurG and OGT. [89, 90] They designed two fluorescently labelled UDP-GlcNAc analogues. They obtained a strong fluorescence polarisation signal with UDP-GlcNAc analogue (Figure 1.10, A) during binding to donor binding site of MurG. The same UDP-GlcNAc analogue was not directly suitable for investigating structurally related OGT. They hypothesised that short linker between the sugar and the fluorophore interfered with binding. Subsequently, they designed a fluorescent probe with longer linker which significantly changed fluorescence polarisation in the presence of OGT. A large commercial compound library was screened against both GTs using individual donor analogues in order to discover small molecular inhibitors. Potential inhibitors were identified from both screenings and then evaluated using a radiometric assay. Based on the same principle, Kiessling's group developed fluorescently labelled UDP analogue (Figure 1.10, B).[91] They screened a diverse compound library against UDP-galactopyranose mutase which uses UDP-sugar donors. Although, the ligand displacement assays are rapid to perform, Walker's results indicated that a particular fluorescently labelled donor is required for each enzyme, even if the natural donor sugar is identical. This results in every GT requiring a differently modified fluorescent donor analogue, which is specific for its catalytic site.

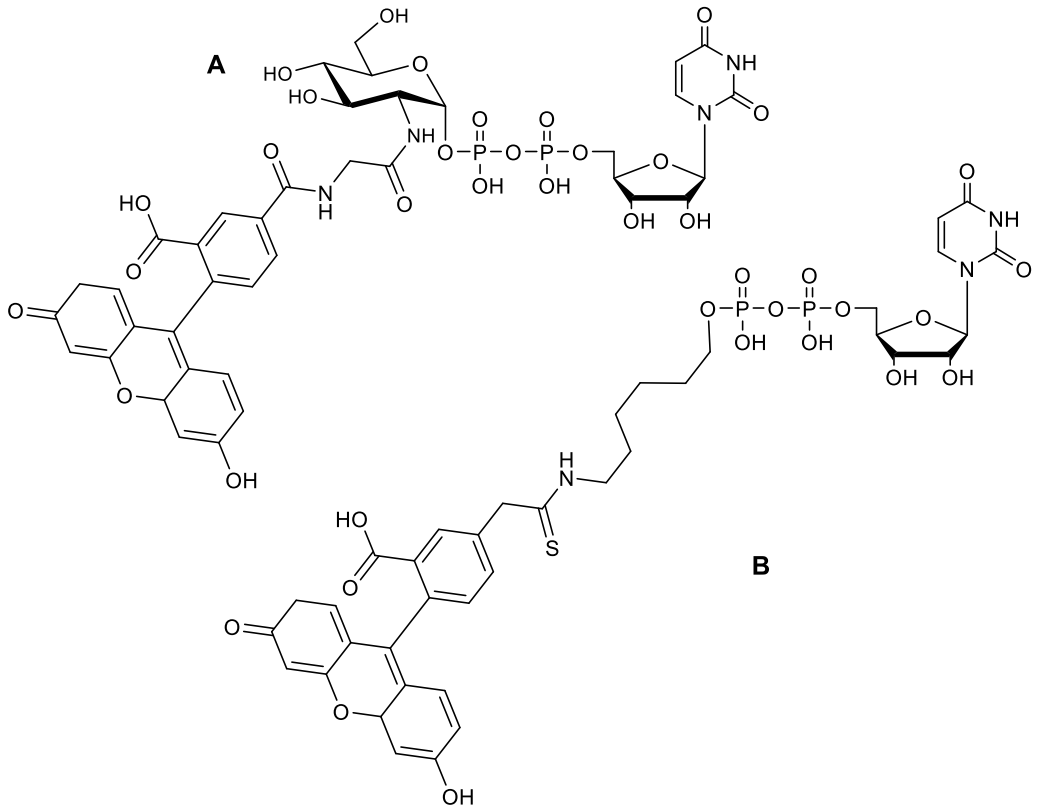


Figure 1.10. Fluorescently labelled donor analogues used in ligand displacement assays for inhibitor screening against Murg by Walker's group[89] (A) and against UDP-galactopyranose mutase by Kiessling's group[91] (B)

Walker's group also monitored *O*-linked glycosylation by a general protease-protection assay where the peptide substrate was labelled with FRET pair and subjected to glycosylation.[56] Protease was added to discriminate between glycosylated and non-glycosylated peptides. The protease cleaved the unglycosylated peptide chain that had a shorter fluorescence signal than the glycosylated peptide chain due to the FRET pair used. The fluorescence signal was, therefore, measured to determine the degree of peptide glycosylation. The assay was validated for HTS in order to identify inhibitors against GTs that catalyse *O*-linked glycosylation.

In recent years, a new type of GT assay has been developed. Wongkongkatep and co-workers investigated binding affinities of an artificial chemosensor.[92] They discovered that the chemosensor selectively bound to pyrophosphate monoesters leading to a significant increase in the fluorescence signal. Therefore, they successfully exploited this approach for the investigation of two galactosyltransferases that produce NDP during the catalytic reaction. This method is carried out in a continuous fashion without labelled substrates and it is applicable for HTS and for many NDP generating GTs. Chandrasekaran and co-workers used a very different approach in their label-free assay. They utilised glycan-modified microspheres as

acceptors for GTs.[93] The glycosylation was monitored with specific fluorescent lectins or antibodies that recognise the newly formed glycoside bonds and an increase in fluorescence was then recorded with a flow cytometry. Although, the method is applicable for HTS, specific glycan microspheres as well as specific lectins or antibodies are required for each GT reaction investigated.

1.3. Glycosidases

Glycosidases, also called glycoside hydrolases, play a diverse role in almost all living organisms, with the exception of some archaeans and some unicellular parasitic eukaryotes.[94] Glycosidases are currently classified into 132 families according to the CAZy database, based on their amino acid sequence similarities.[17, 18] The families can be further combined into clans sharing common tertiary structures and evolutionary origin.[95] About three quarters of the glycosidase families have structural information available, whereas, about one-third of the GT families have a known structure.[96] Glycosidases are more stable enzymes than GTs and they are easily isolated from numerous natural sources such as micro-organisms and plants. Glycosidases catalyse the cleavage of the glycosidic bonds in glycosides and glycans.[97] The hydrolysis of the glycosidic bond is achieved by glycosidases with overall inversion or retention of anomeric carbon configuration of the substrate. Glycosidases can be classified into endo- and exo-acting enzymes. Endo-acting glycosidases hydrolyse the glycosidic bond in the interior of the saccharide chain, whereas exo-acting glycosidases act at the nonreducing end, producing a monosaccharide as a leaving group.[98]

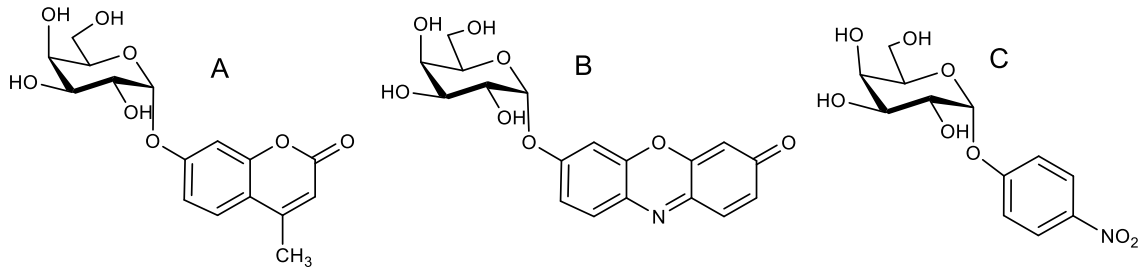


Figure 1.11. Chromogenic substrates for the analysis of α-galactosidase. A) 4-methylumbelliferyl-α-D-galactopyranoside, B) resorufinyl-α-D-galactopyranoside and C) *p*-nitrophenyl-α-D-galactopyranoside

Numerous assays have been developed for monitoring glycosidase activity and for inhibitor screening.[94] Glycosidase assay development is simpler than for GTs because the hydrolysis reaction occurs in a single substrate manner. Generally, natural glycosidase substrates are poly- or oligosaccharides and do not contain chromophores, therefore, chromogenic substrates have been developed for monitoring the activity of various glycosidases. The most recent fluorogenic substrate, resorufinyl-α-D-galactopyranoside (Figure 1.11, B), was used in a HTS assay in order to discover inhibitors or activators against α-galactosidase.[99] The α-galactosidase reaction with resorufinyl-α-D-galactopyranoside produces galactose and resorufin, which is strongly fluorescent with an excitation at 573 nm and an emission at 590 nm. Chromogenic assays that uses 4-methylumbelliferyl-α-D-galactopyranoside[100] or *p*-nitrophenyl-α-D-galactopyranoside[101] as substrates (Figure 1.11, A and C) are available for monitoring activity of α-galactosidase.

1.4. Project objectives

GTs are potential drug targets as they are involved in essential biological processes, such as in cell adhesion and in bacterial cell wall biosynthesis. Currently, there are only a handful of effective GT inhibitors that could be used as tools for studying the mechanism of GTs and the biological function. In order to discover novel GT inhibitors, robust and operationally simple screening assays are required. Therefore, the key objective of this project was to develop GT assays for the discovery of novel inhibitors and for inhibitor characterisation.

- Previous studies showed that a strongly fluorescent novel sugar nucleotide analogue (5-(5-formyl-2-thiophene)-UDP-Gal, 5FTUDP-Gal, Figure 1.12) was a potential tool to investigate GTs. The preliminary investigations demonstrated that the UDP-Gal analogue bound to several galactosyltransferases, which resulted in a fluorescence intensity quenching and this read-out was used for the design of a novel ligand displacement assay (Figure 1.12).[102] Therefore, the aim was to optimise and validate this assay set up for inhibitor screening, and also to miniaturise the ligand displacement assay for 384 well microplates (Chapter 2).

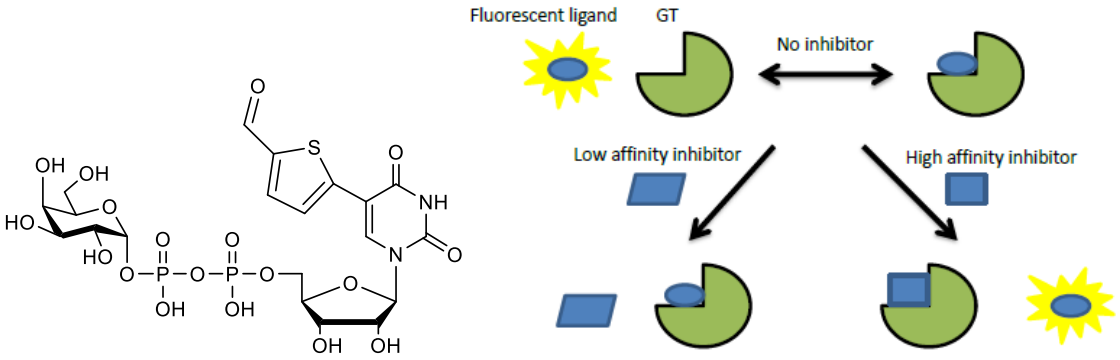


Figure 1.12. Fluorescent UDP-Gal analogue (5-(5-formyl-2-thiophene)-UDP-Gal) and principle of the ligand displacement assay

- The fluorescent UDP-Gal analogue was previously successfully used in the ligand displacement assay with α1,4-galactosyltransferase C from *N. meningitidis* (LgtC), α1,3-GalT from *Bos taurus* (*B. taurus*) and from *Homo sapiens (H. sapiens)*, α1,3-GalT mutant (AA(Gly)B).[102] Therefore, the aim was to investigate the applicability of the fluorescent UDP-Gal analogue to other glycosyltransferases by using the ligand displacement assay (Chapter 2).

- The aim was to develop a novel biochemical assay in order to investigate activities of GTs and for characterisation of potential GT inhibitors. (Chapter 3). The novel GT assay based on an unnatural fluorescent acceptor and coupling of the LgtC reaction to a glycosidase reaction. The aim was also to develop a HPLC-based method in order to investigate activities of GTs and to run the HPLC-based method in parallel during the novel assay development. The development of the novel GT assay was not achieved. Therefore, the aim was to optimise and validate an existing phosphatase coupled biochemical assay to monitor GT catalysed reaction and to characterise potential GT inhibitors.

- In order to identify novel LgtC inhibitors, the aim was to use the optimised and validated ligand displacement assay for compound library screening (Chapter 4). The aim was to screen two different compound libraries and screen the second library with the ligand displacement assay and with the phosphatase coupled assay in order to investigate differences of the assays by comparing the outcomes. The aim was also to characterise the hits identified from both screenings by the phosphatase coupled biochemical assay.

2. Fluorescence-based ligand displacement assay

2.1. Introduction

Generally, glycosyltransferase (GT) assays that have previously been developed are not applicable for high-throughput screening (HTS). Radiochemical assays are sensitive and widely used for monitoring GT activity, however, the disadvantages are the availability of radiolabeled substrates, washing steps and removal of unreacted radiolabeled substrates from the labelled products which often requires chromatographic methods.[66] In general, assay formats that include chromatographic separations are unsuitable for compound library screening due to their time-consuming and non-automated protocols. In addition, chromatographic methods such as HPLC and CE are commonly coupled with UV-detectors which limit monitoring of the primary product of GT reaction if substrates are not labelled with a chromophore.[63] However, some GT assays that are suitable for HTS have been reported,[63, 66] although some of these methods have some considerable limitations. For example, immunological assays[103] require specific antibodies for each reaction product and in coupled enzyme assays,[84] large quantities of different enzymes are utilised which makes the assay expensive. Many of the high-throughput methods have been developed for monitoring activity,[84] substrate specificity[92] and saturated mutagenesis[82] of GTs, but few of these assays have been applied for compound library screening.

A novel fluorescent UDP-Gal analogue has been previously developed[102] and preliminary studies demonstrated that this fluorophore binds to various retaining GalTs (LgtC, α1,3-GalT from *B. Taurus* and from *H. Sapiens*) in a similar orientation to the natural substrate. Fluorescence intensity quenched upon the fluorophore binding to the catalytic pocket and the fluorescence was restored in the presence of a high affinity binder.[102] This combination was then used for the design of a novel ligand displacement assay (LDA). A small compound library was screened suggesting that the fluorophore could be used as a tool for the discovery of galactosyltransferase (GalT) inhibitors.[104]

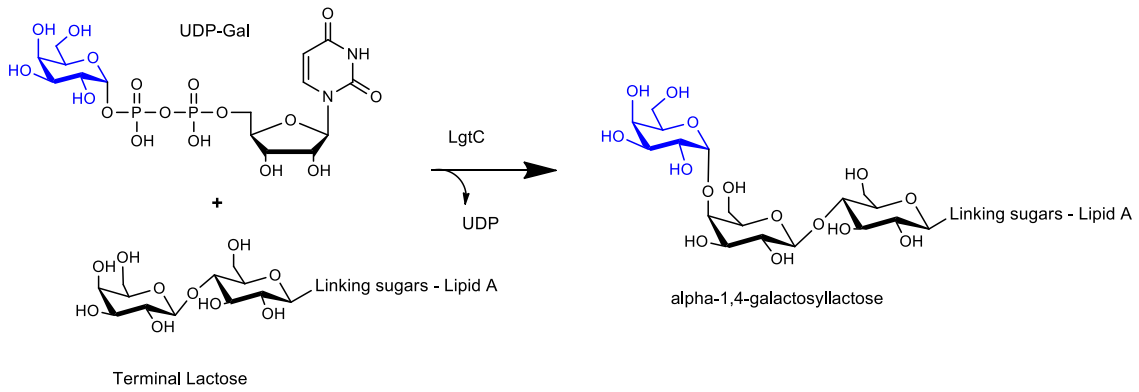
In principle, the assay format can be easily adapted for HTS and possibly applied to other sugar-nucleotide-dependent enzymes. Therefore, in order to use the novel LDA for inhibitor screening in HTS format, an optimisation process was required. The preliminary results indicated that the fluorophore could be used with several GTs, and this encouraged

investigating the use of a fluorescent UDP-Glc analogue with two glucosyltransferases. Therefore, two other bacterial GTs, TcdB from *Clostridium difficile* (*C. difficile*) and NGT from *Actinobacillus pleuropneumoniae* (*A. pleuropneumoniae*) were investigated. A bacterial GalT, α1,4-galactosyltransferase C (LgtC) was used as a model enzyme for the LDA optimisation and validation studies because of the possibility to express large quantities of this enzyme and the preliminary studies demonstrated that the UDP-Gal analogue bind to LgtC. In addition, previous studies demonstrated that the UDP-Gal analogue binds to mammalian GT, β1,4-GalT and behaves as a substrate.[102] The determined K_m of UDP-Gal analogue was 74 μM whereas, K_m of the natural substrate, UDP-Gal, was 46 μM. Therefore, β1,4-GalT was introduced to the studies as a positive control with LgtC.

Prior to introducing the ligand displacement assay, and the assay optimisation process in more detail, the GTs utilised in this chapter (LgtC, TcdB, NGT and β1,4-GalT) will be introduced. The protein structure and function of the donor substrate binding of the studied GTs must be understood so that the binding mode of the fluorophore and experimental results can be analysed.

2.1.1. LgtC

Lipopolysaccharide α1,4-galactosyltransferase C (LgtC) from *N. meningitidis* is a retaining galactosyltransferase and it belongs to the GT family 8.[16] LgtC is involved in the biosynthesis of meningococcal LOS by transferring α-D-galactose from UDP-Gal to lactose of the terminal lactose moiety (Scheme 2.1).[16] *N. meningitidis* is a Gram-negative pathogenic bacteria that is only found in humans and causes a life-threatening meningitis and meningococcemia worldwide.[50] Bacterial meningitis is generally treated with antibiotics although drug-resistant bacteria are increasingly problematic, making widely used antibiotics ineffective.[105] An effective vaccination exists for meningococcal serogroups A, C, Y and W135, and very recently a novel vaccine against serogroup B has been discovered.[106] However, the protection against these infections is 73 to 87 % caused by the serogroup B, [107] therefore, new antibiotics to treat meningococcal infections are still needed.



Scheme 2.1. LgtC catalysed reaction

The high-resolution crystal structure of LgtC has been determined in complex with Mn^{2+} and UDP-2-deoxy-2-fluoro-galactose (UDP-2FGal), a sugar donor analogue, in the presence and in the absence of the acceptor sugar analogue, 4'-deoxylactose.[16] The structure of LgtC is formed from nine β-sheets and fourteen α-helixes that are organised into two domains (Figure 2.1). The large N-terminal domain contains the active site that is surrounded by β-sheets. LgtC has four Asp-x-Asp (DXD, x is any amino acid) motifs and two of them are located at the active site. The C-terminal domain is smaller and has membrane associated functions.[16] Upon donor substrate binding to LgtC, the enzyme undergoes significant structural changes and two flexible loops are proposed to become ordered, from open (inactive) to closed (active) conformation (Figure 2.1, flexible loops in green).[16, 108] However, a substrate free crystal structure of LgtC is still unavailable and therefore, the position of the flexible loops is unclear in the *apo* stage.

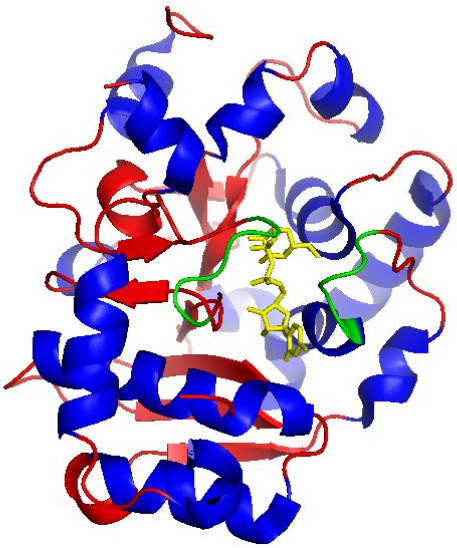


Figure 2.1. Crystal structure of LgtC complex with UDP-2FGal in yellow (ball- and- stick model) Flexible loops are coloured in green, β-sheets in red and α-helixes are in blue (PDB: 1G9R)

After the three-dimensional structure was published, mechanistic studies of LgtC have been carried out experimentally and computationally.[27, 28, 108, 109] Kinetic analysis showed that LgtC follow an ordered bi-bi kinetic mechanism where UDP-Gal binds first to the active site, followed by the acceptor.[109] Recent investigations proposed that LgtC uses a S_Ni-like reaction, where transfer occurs via a short lived oxocarbonium-phosphate ion pair intermediate or via an oxocarbenium ion-like transition state.[28, 108]

LgtC has been found to be involved in the biosynthesis of the lipopolysaccharide (LPS) of human pathogen *Haemophilus influenzae* (*H. influenzae*).[110] Phase-variable genes, including lgtC, express the galabiose motif (Gal-α1,4-Galβ) that is recognised by monoclonal antibody 4C4[111] and autotransporter protein Hap[112] in *H. influenzae*. Inactivation of lgtC has been shown to lead to decreasing the virulence of *H. influenzae* in an infant rat model of invasive infection.[110] The lack of Hap on the bacterial cell surface leads to a decreasing interaction with the host.[112] *H. influenzae* R2866 strain (lgtC "on"), isolated from a meningitis patient and its derivative R3392 (lgtC "off") strain have been used to study the role of lgtC in virulence of *H. influenzae*[113-115] and the results strongly indicate that active lgtC increases virulence of *H. Influenzae*. Because LgtC has shown to play an important role in the biosynthesis of bacterial cell wall, it is of great interest for the future design of new type of inhibitors that could block the biosynthesis of the negative bacterial LOS.

2.1.2. β1,4-GalT

Golgi resident type II membrane protein β1,4-Galactosyltransferase I (β1,4-GalT) is an extensively studied GT. To date, β1,4-GalT has been reported as involved in two different enzymatic activities. β1,4-GalT synthesises a disaccharide unit, Gal-β1,4-GlcNAc (*N*-acetyllactosamine) by transferring galactose from UDP-Gal to GlcNAc and therefore continuing the oligosaccharide chain elongation on the cell surface.[21] In the presence of mammary gland-specific protein α-lactalbumin, β1,4-GalT forms a lactose synthase complex. Binding of α-lactalbumin changes the acceptor specificity of β1,4-GalT and galactose is transferred to glucose, yielding lactose (Gal-β1,4-Glc), a milk sugar.[21, 116-118] The diversity of β1,4-GalT was recently expanded when a poly-LacNAc moiety of the branched β1,6-GlcNAc was found binding to a new binding region of β1,4-GalT.[119]

Since the first crystal structure of β1,4-GalT was published its biochemical and kinetic mechanisms have been widely studied. The *N*-terminus of the active domain is a Rossmann-like fold which recognises donor substrates while the C-terminal assists acceptor binding.[120] The catalytic pocket is located between the two domains and two flexible loops, a small (residues 313-316) and a long (residues 345-365) loop assists donor binding. During the catalytic cycle, the conformation of β1,4-GalT changes radically. Mn^{2+} binds first to the active site while the enzyme remains in an open conformation. During the sugar donor binding two flexible loops undergo significant conformational changes creating the acceptor binding site (Figure 2.2).[117, 121] Tryptophan 314 (W314) plays a crucial role in donor binding and β1,4-GalT activity. During the closed conformation (active) the W314 residue is inside the catalytic pocket, interacting with both the UDP-Gal and the acceptor molecule.[122]

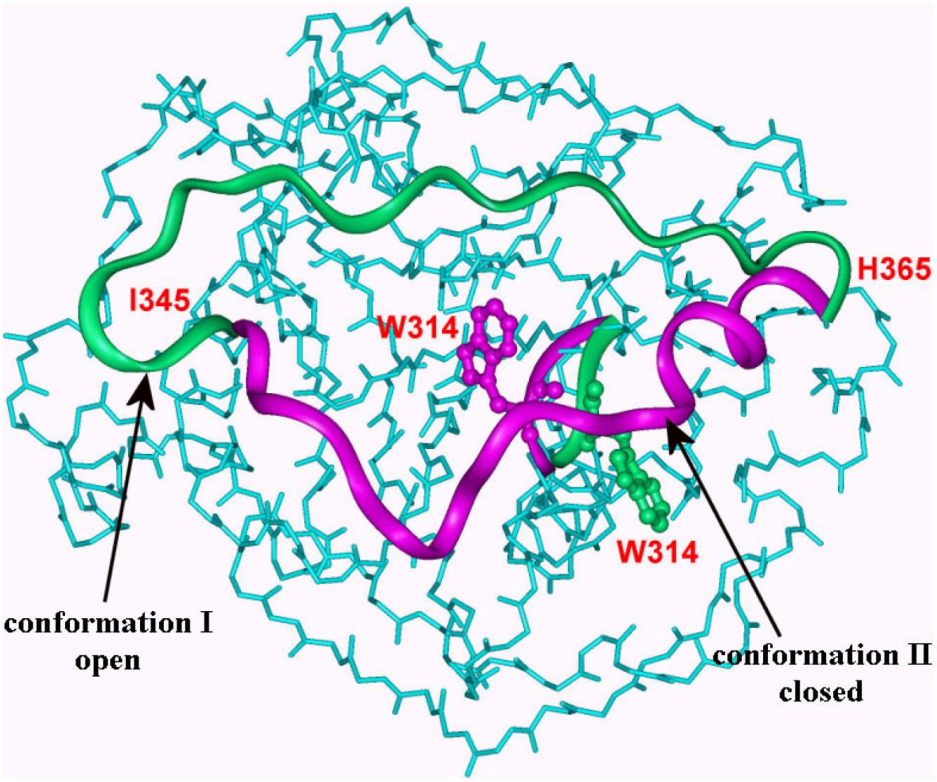


Figure 2.2. A significant structural change in open (Green, conf. I) and closed (Magenta, conf. II) conformations of β1,4-GalT. The long loop (residues I345-H365) and the short loop containing W314 (ball-and-stick model). W314 is facing in to the catalytic pocket in closed conformation (active conformation). (Taken from reference[121])

β1,4-GalT is expressed in all mammalian tissues. On the eukaryotic cell surface it can function as a receptor molecule and participate in cellular interactions by binding to particular extracellular glycoside substrates. Therefore, β1,4-GalT has an important role in several cellular functions including neurite extension,[123] sperm-egg interaction,[124] cell adhesion and migration.[125] β1,4-GalT producing gene knockout in mice resulted in decreased β1,4-GalT activity in most tissues[124] and the mice fertility was decreased,[126] differentiation of epithelial cells, and increased lethality and growth retardation.[127] Recent studies have suggested that in humans, β1,4-GalT is involved in growth and adhesion of cancer cells[128, 129] and may play a role in embryo implantation.[130]

2.1.3. TcdB

Bacterial toxin (TcdB or Toxin B), released by *C. difficile* causes severe antibiotic-associated diarrhea and pseudomembranous colitis.[131] TcdB targets GTPases, proteins that are involved in multiple eukaryotic cellular functions including signalling, cell cycle progression, actin cytoskeleton regulation, gene transcription and various enzyme activities. Rho, Rac and Cdc42 belong to GTPases[132, 133] and TcdB inactivates them by glucosylation using UDP-Glc as a donor substrate.[134, 135] TcdB is a large clostridium toxin and it is formed from four domains that are related to cellular uptake and toxicity. The C-terminal binding domain identifies target cell surface proteins and assists in binding.[133] The translocation domain uptakes the two remaining domains to the cytoplasm by forming a transmembrane pore.[136, 137] The activity of the cysteine protease domain (CPD), the fourth domain, is responsible for the cleavage of glucosyltransferase domain that is responsible for the toxicity of *C. difficile*. Only the N-terminal segment of the CPD is released to the cytosol where the GTPases can be glucosylated causing cell apoptosis.[138, 139]

The crystal structure of the glucosyltransferase domain of TcdB has been reported in the presence of UDP-Glc and a manganese ion.[134] The crystal structure model of the active domain contains 543 residues and encloses chiefly parallel β-sheets surrounded by sheet segments. This topology closely resembles the structure of LgtC and TcdB is assigned to GT type A fold.[134] The TcdB catalysed reaction retains the anomeric conformation and is dependent on Mn^{2+}, in addition, two flexible loops are proposed as involved in both substrate binding and in product release.[134]

As the acceptors of TcdB are generally unavailable, the LDA could be an alternative method for the inhibitor discovery of this bacterial toxin because acceptors are not required in the assay format. In order to expand the LDA for glucosyltransferases, based on the structural similarities and the mechanism similarities with LgtC (both GT-A fold and retaining GT) it was assumed that the fluorescent UDP-Glc analogue binds to TcdB similarly to the fluorescent UDP-Gal analogue binds to LgtC, resulting in a decrease of the fluorescent signal.

2.1.4. NGT

N-glycosyltransferase (NGT) from *A. pleuropneumoniae*, a homologue of HMW1C from *H. influenzae*, was recently discovered. The NGT is an inverting GT and belongs to the GT-41 family.[140-142] NGT and HMW1C modify the asparagine residues of the HMW1 adhesion protein from *H. influenzae* by forming N-linkages with glucose or galactose and they are also able to create bonds between two hexoses.[143] N-glycosylation of HMW1 protein leads to bacteria adhesion to the respiratory epithelial cells, therefore, glycosylation is a significant step in the pathogenesis of *H. Influenzae.*[140]

Recently, the crystal structure of NGT has been reported in the presence and the absence of UDP-Glc and in the presence of a known glycopeptide.[141] NGT consists of three domains, the N-terminus an all α-helical domain, and the GT-B fold which is divided into two domains, GT-1 and GT-2 on the C-terminus. The core structures of GT-1 and GT-2 are similar and the UDP-hexose binding site is positioned at their interface. Binding of UDP includes slight conformational changes in the NGT structure.[141] The side chains and the three loops on the UDP-hexose's binding site co-operate by placing the amino acid near UDP and assist the formation of the hydrogen bonds during sugar donor binding.[141]

Similar to TcdB, the LDA was an attractive method for NGT because the acceptor proteins of NGT are not easily available. In addition, the LDA has been investigated only with retaining, GT-A fold galactosyltransferases. Therefore, NGT which possesses a GT-B fold and which is an inverting glucosyltransferase was of interest to study the applicability of the LDA.

2.2. Fluorescence-based displacement assay

A fluorescent UDP-Gal analogue (5-(5-formyl-2-thiophene)-UDP-Gal, 5FTUDP-Gal, Figure 2.3) was recently characterised[102] and the results demonstrated a slow turnover when 5FTUDP-Gal was utilised as a substrate and therefore 5FTUDP-Gal was unsuitable for the development of biochemical assays. The preliminary studies illustrated that the fluorescence intensity was quenched when the 5FTUDP-Gal was incubated in the presence of various GalTs including LgtC and α1,3-GalT, and the binding occurred in a similar orientation to the natural substrate UDP-Gal.[102] The protocol of the assay was of interest due to its simplicity and aptitude for HTS. Therefore, after sufficient assay optimisation, the aim was to apply the LDA to a compound library screen against GTs. Additionally, to enable more financially viable assay, the LDA adaptation to 384 well microplate format was investigated.

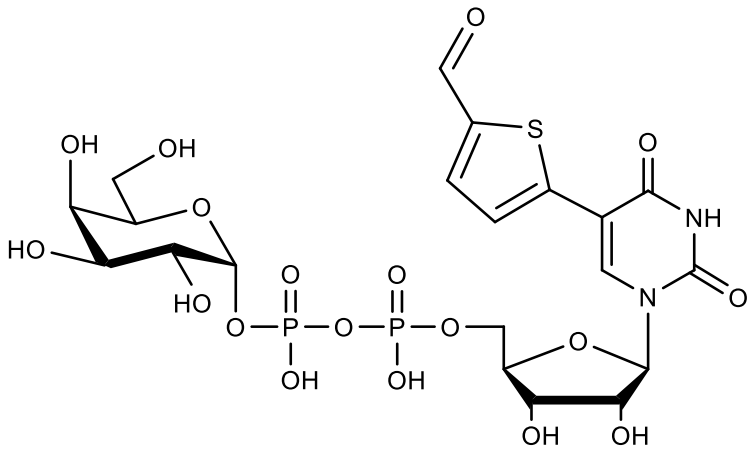


Figure 2.3. Structure of the fluorophore used in the LDA, 5-(5-formyl-2-thiophene)-UDP-Gal

The fluorescence based ligand displacement assay can be divided into two sets: ligand binding and ligand displacement. Before performing the ligand displacement experiments, the ligand must bind to the donor binding site of desired enzyme. The ligand binding can be examined by incubating the ligand (5FTUDP-Gal, fluorophore) in various enzyme (*e.g.* LgtC) concentrations. If the ligand binds to the enzyme, then the fluorescence quenches with increasing concentration of the enzyme (Figure 2.4 A). If fluorescence quenching is not observed, then the particular enzyme is unsuitable for ligand displacement experiments. Fluorescence can be restored by high affinity binders (*e.g.* natural substrates, UDP-Gal or UDP) that release the ligand from the catalytic pocket of the enzyme and the increase fluorescence intensity can be monitored (Figure 2.4 B).

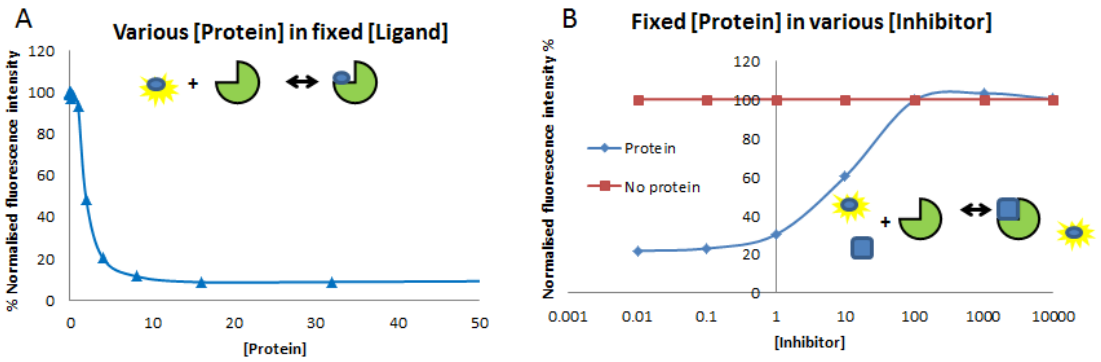


Figure 2.4. Principle of the ligand displacement assay. A) Fluorescence intensity quenches when the ligand (5FTUDP-Gal) is incubated with increasing concentration of a protein (*e.g.* LgtC). B) Fluorescence intensity is restored with increasing concentration of a high affinity binder or inhibitor (*e.g.* UDP)

The preliminary studies of the LDA were performed with bovine α1,3-GalT as a model enzyme.[102] The control experiments included incubations in the absence and in the presence of $MnCl_2$ and various concentrations of $MnCl_2$, and in the absence and presence of α1,3-GalT. These observations indicated that the fluorescence quenching was due the fluorophore binding. The binding site of the fluorophore was investigated with the natural donor, UDP-Gal, which displaced the fluorophore from the active site of α1,3-GalT. Additionally, the control experiments included an incubation time and stability of the fluorophore. The LDA was inadequate for compound library screening and the assay required more control experiments regarding the functionality, the stability and the reproducibility. Therefore, the following investigations on the ligand binding affinity and the ligand displacement, using LgtC as a model enzyme were performed:

- Reaction volume
- Concentration of the fluorophore (5FTUDP-Gal)
- Concentration of DMSO
- Incubation time
- Buffer and buffer pH
- Concentration and activation of LgtC
- Concentration of $MnCl_2$
- Non-specific binding
- Reproducibility
- Assay miniaturisation to 384 well microplate format

2.2.1. Investigating: reaction volume, concentration of fluorophore and DMSO, buffer and incubation time

The preliminary studies were performed in 50 mM Tris/HCl (tris (hydroxymethyl)-aminomethane/hydrochloric acid) buffer including 10 mM $MnCl_2$ at pH 7.0, reaction volume was 200 µL, fluorophore concentration 200 nM, incubation time 10 minutes and reaction temperature 30 °C.[102] The lower reaction volume was investigated in order to keep the material consumption as low as possible during the assay optimisation. The impact of the reaction volume was investigated by performing a calibration curve of the fluorophore (5FTUDP-Gal) at 100 µL and 200 µL. Linearity was identical with increasing concentration of the fluorophore (Figure 2.5) and therefore, the reaction volume was reduced to 100 µL for all subsequent experiments. The previously used fluorophore concentration (200 nM) was in the linear range, therefore, it was unchanged.

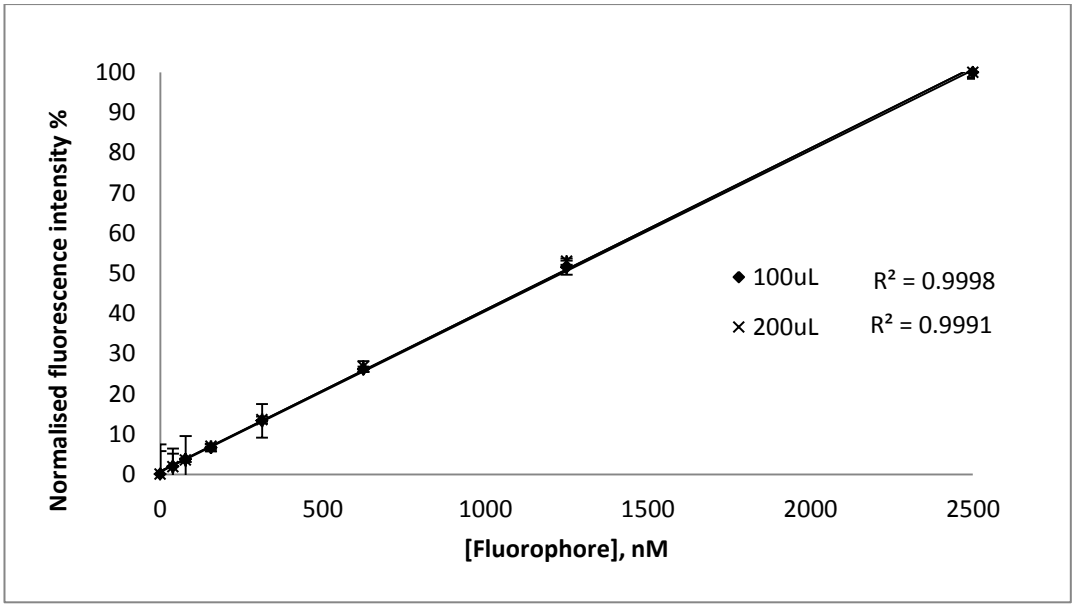


Figure 2.5. Linearity of fluorophore in the range of 39 nM to 2500 nM at total volume of 100 µL and 200 µL

Tris/HCL buffer (pKa 8.06 at 25 °C[144]) has poor buffering capacity below pH 7.5[144, 145] and in order to ensure that the pH remained stable during enzymatic reactions, Tris/HCl buffer was compared with HEPES (*N*-(2-Hydroxyethyl)-piperazine -*N*'-ethanesulfonic acid) in various pH. HEPES has a lower pK_a (pKa 7.48 at 25 °C[144]) and its effective pH range is slightly lower: 6.8-8.2. In addition, HEPES has a low metal-binding constant. Therefore, it is particularly suited to investigate metal-dependent enzymes.[146]

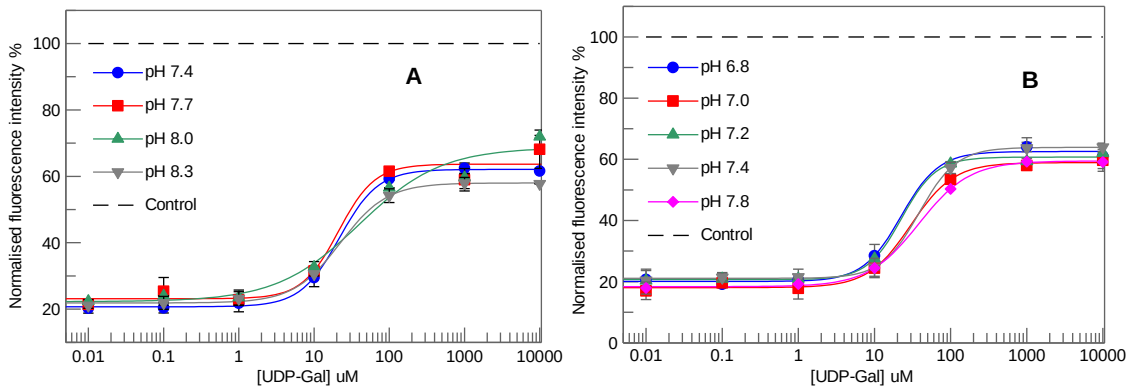


Figure 2.6. Ligand displacement assay with various UDP-Gal concentrations in Tris/ HCl (A) and in HEPES (B) buffers in various pH

The LDA was adapted to investigate differences between Tris/HCl and HEPES buffer at various pH. The assay functioned over the entire pH range tested with Tris/HCl and HEPES buffers (Figure 2.6), indicating that fluorophore binds to LgtC and can be replaced with high affinity binder, UDP-Gal. The specified IC_{50} values (Table 2.1) suggested that the fluorophore binds more tightly to LgtC with HEPES buffer at pH 7.4 and above. Above pH 7.4 more UDP-Gal is required to replace the fluorophore. However, 13mM HEPES buffer at pH 7.0 was chosen for subsequent experiments, because pH 7.0 is widely used in various GT assays[82, 84] and the LDA functions well at this pH.

Table 2.1. IC_{50} for UDP-Gal in Tris/HCl and HEPES buffers at various pH

pH (TRIS/ HCl)	IC_{50}, µM (TRIS/ HCl)	pH (HEPES)	IC_{50}, µM (HEPES)
7.4	22 ± 2	**6.8**	22 ± 8
7.7	21 ± 10	**7.0**	30 ± 4
8.0	40 ± 23	**7.2**	23 ± 3
8.3	22 ± 2	**7.4**	39 ± 2
		7.8	38 ± 2

Pre-incubation and incubation times were investigated in the LDA. The fluorophore binding and displacement is assumed to occur rapidly, therefore, a long incubation time is not necessary. To support this hypothesis, a control experiment was performed with various pre-incubations. The final reading was taken 20 minutes after the last component was added.

- Incubation 1: Binder (UDP in this case) and fluorophore incubated for 10 minutes then LgtC added
- Incubation 2: Binder and LgtC incubated for 10 minutes then fluorophore added
- Incubation 3: LgtC and fluorophore incubated for 10 minutes then binder added

The results indicated that displacement occurs similarly in all cases because the IC_{50} values did not change significantly (Table 2.2). For practical reasons, the fluorophore and the binder or an inhibitor of interest were incubated first and LgtC was added last (incubation 1). This was the chosen pipetting order for all subsequent experiments.

Table 2.2. IC_{50} values in different incubation modes and in various final incubation times. Experiments performed at 30 °C

Incubation mode[a]	IC_{50} (UDP), µM	Incubation time, min	IC_{50} (UDP-Gal), µM	IC_{50} (UDP), µM
1	24 ± 3	10	26 ± 2	35 ± 2
2	18 ± 3	20	31 ± 2	36 ± 2
3	26 ± 3	30	35 ± 2	34 ± 1
		60	39 ± 3	39 ± 2

[a]Reaction mixtures incubated 20 minutes before final reading was taken

The final incubation time, the incubation time after the final component (LgtC) addition, was studied by displacing the fluorophore with UDP-Gal or UDP. The fluorophore displacement can be monitored in real time. The results demonstrated that UDP and UDP-Gal displace the fluorophore rapidly as the incubation time did not have significant effect on the IC_{50} value (Figure 2.7 and Table 2.2). The fluorophore possibly photobleaches slightly during the incubation, therefore, increased IC_{50} values were observed with increasing incubation time. 20 minutes incubation time was chosen for all future experiments because the balance of the displacement was reached and photobleaching of the fluorophore can be avoided. In addition, 20 minutes incubation time provided a longer monitoring window in compound screening if potential inhibitors displace the fluorophore slowly.

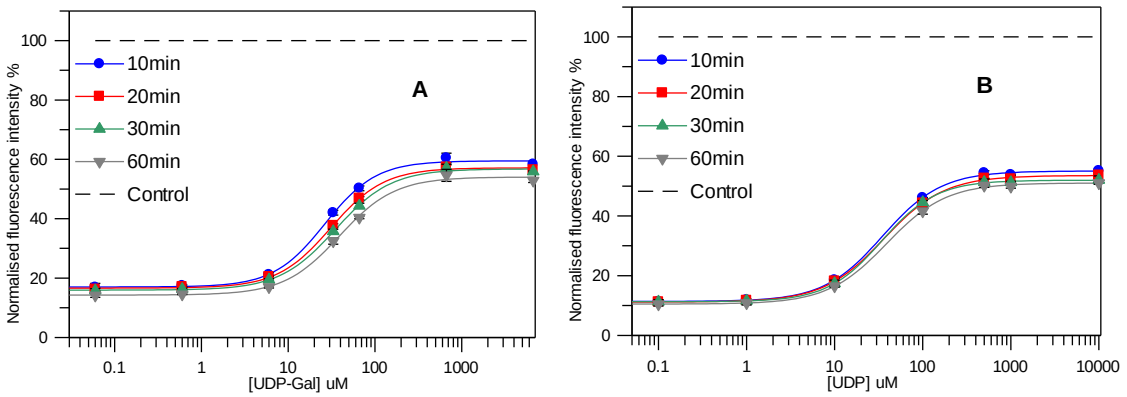


Figure 2.7. Displacement assay with UDP-Gal (A) and UDP (B) with various incubation times

Finally, it was important to ensure that the ligand displacement assay can tolerate DMSO. DMSO increases the solubility of organic compounds which are often poorly soluble in aqueous buffers. Therefore, DMSO is a necessary additive in assay buffer when the assay is designed for compound screening. The effect of DMSO on the LDA was investigated by varying the concentration from 0 % to 10 % (v/v). The results demonstrated that the ligand displacement assay tolerated DMSO up to 10 %. DMSO did not have a significant effect on the fluorescence intensity signal or the fluorophore displacement either with UDP-Gal or UDP (Figure 2.8).

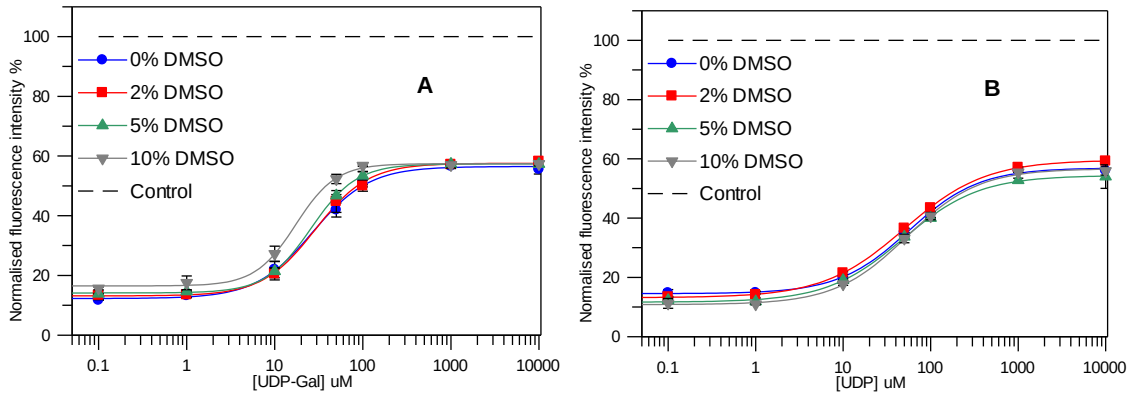


Figure 2.8. Effect of DMSO on the LDA with various concentrations of UDP-Gal (A) and UDP (B)

The control experiments indicated that the LDA is not effected by the assay buffer, the type of incubation used or high DMSO concentration, which makes the LDA very robust. All the experiments above were performed at 30 °C by incubating the fluorophore with binder prior to the enzyme addition, total reaction volume of 100 µL, 200 nM of the fluorophore and all reactions were performed in 13 mM HEPES (pH 7.0) buffer. These conditions were kept in all subsequent experiments unless otherwise indicated.

2.2.2. Influence of LgtC activation

Kinetic studies have shown that LgtC requires activation with DTT (dithiothreitol) to reach full enzymatic activity.[147] However, the crystal structure studies of LgtC suggested that there are no cysteine (Cys) residues that react in the donor substrate binding site.[16] During the acceptor binding a conformational change occurs and Cys 246, located in the flexible loops, forms a hydrogen bond with one of the oxygen atoms of lactose. These findings suggest that LgtC activation affects only the acceptor binding and is not necessary if donor binding is monitored.

However, there is no three-dimensional data available for LgtC in complex with the fluorophore and it is currently unknown if there is more than one fluorophore molecule bound to the donor binding site or if the fluorophore is bound partly to the acceptor binding site. Therefore, a control experiment was carried out in order to investigate whether the activation changes the required LgtC concentration to obtain a fluorophore binding curve by assuming that more fluorophore molecules are bound to the activated LgtC. Effect of activation was investigated by comparing activated and non-activated LgtC in the binding affinity assay.

The degree of fluorescence quenching was identical when the binding curves of activated and non-activated LgtC were compared (Figure 2.9). The results suggested that activation did not affect the fluorophore binding and confirmed the hypothesis that the fluorophore probably does not bind to the acceptor binding site. Therefore, activation may only have an effect when the acceptor binds to the LgtC-donor complex leading to the formation of reaction products. In this assay mode, the acceptors are not utilised and binding of the fluorophore is monitored, therefore, all subsequent experiments were performed with non-activated LgtC.

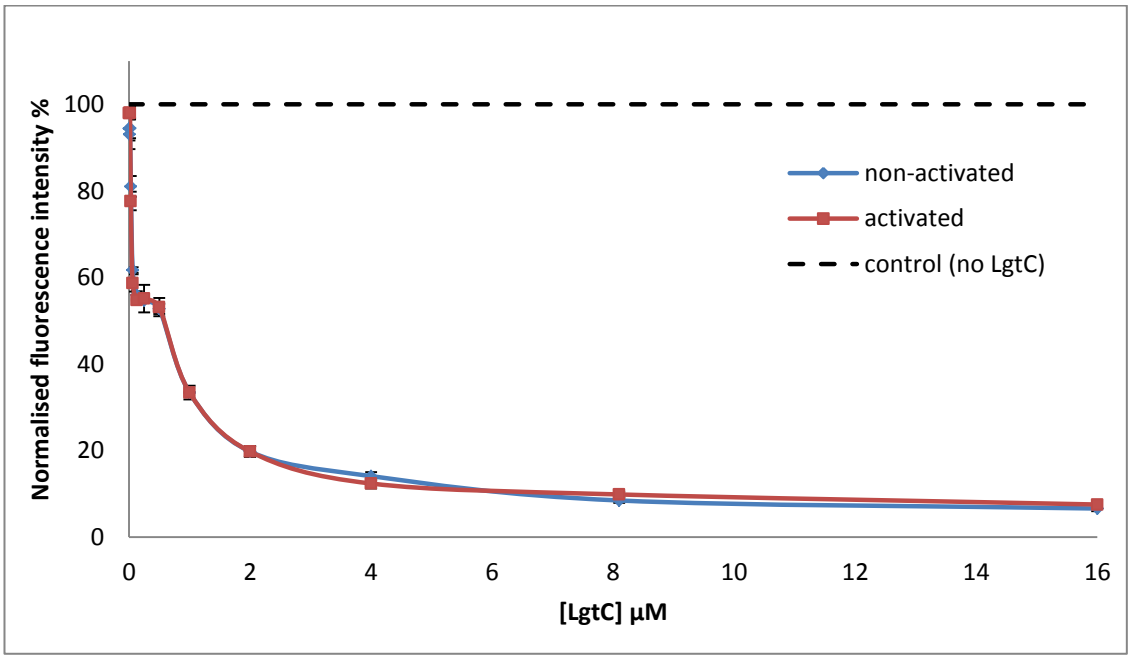


Figure 2.9. Effect of LgtC activation with DTT compared with non-activated LgtC in the binding affinity assay

2.2.3. Influence of metal

The manganese (II) ion is required for the activity of LgtC as it binds to the DXD motif and assists UDP-Gal binding to the active site.[16] Therefore, it is possible that, as with the natural substrate, the fluorophore binding will occur only in the presence of Mn^{2+}. The preliminary studies with α1,3-GalT demonstrated that fluorescence quenching was not observed in the absence of Mn^{2+} and significant quenching occurred only when Mn^{2+} was present.[102] A control experiment was performed to examine the requirement of Mn^{2+} on the LDA with LgtC in the presence Mn^{2+} and without Mn^{2+} addition.

Surprisingly, different results were obtained with LgtC, compared to α1,3-GalT. Fluorescence was quenched in both, in the presence and without addition of Mn^{2+} (Figure 2.10). The reason for the quenching was identified from the LgtC expression and purification protocol (Chapter 2.7.1). The LgtC stock solution contains 5 mM of Mn^{2+} and this concentration was enough to create fluorophore binding to LgtC.

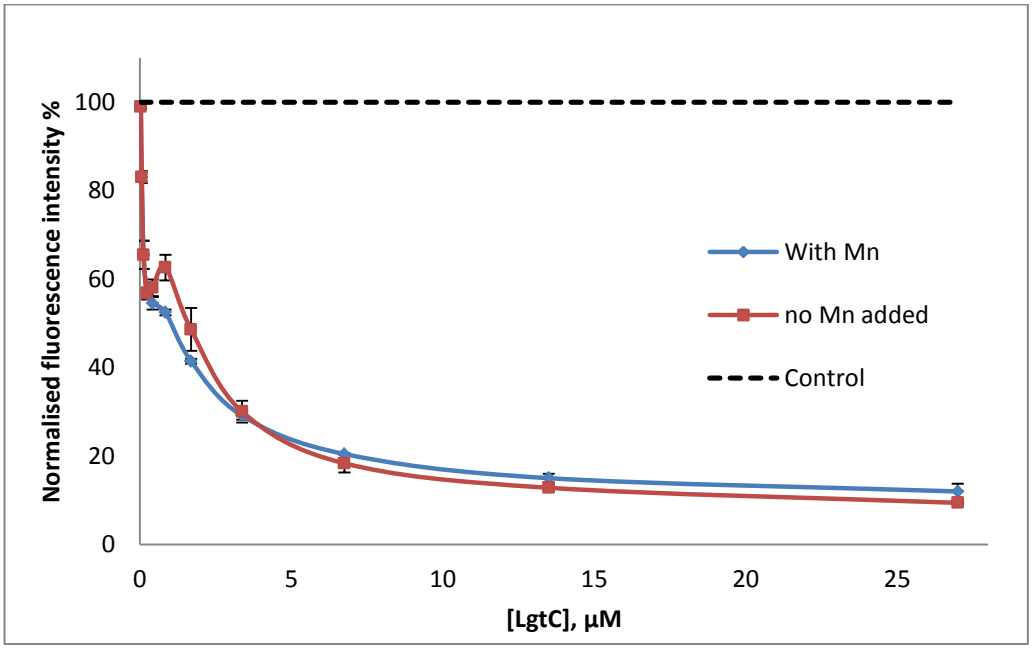


Figure 2.10. Effect of manganese for fluorophore binding to LgtC. Fluorophore binding affinity monitored with increasing concentration of LgtC without Mn^{2+} addition and in the presence of 10 mM of Mn^{2+}

The impact of the Mn^{2+} was studied with the LDA without addition of Mn^{2+} (10 µM of Mn^{2+} from LgtC stock) and in the excess[104] (10 mM) of Mn^{2+}. The results indicated that Mn^{2+} has a significant effect on IC_{50} values. IC_{50} increased from 32 µM to 213 µM with UDP-Gal and 30 µM to > 1000 µM with UDP (Figure 2.11) in the presence of 10 mM Mn^{2+}. These results suggested that more fluorophore is bound to LgtC or the fluorophore binds very tightly to LgtC in high Mn^{2+} concentration. Recently Chan and co-workers studied conformational dynamics of LgtC and they showed that donor sugar analogue (UDP-2FGal) associated with LgtC only in the presence of a metal ion.[108] Therefore, it was assumed that at high Mn^{2+} concentration interaction between the fluorophore and LgtC is increased and thus higher concentration of binder is required to displace the fluorophore.

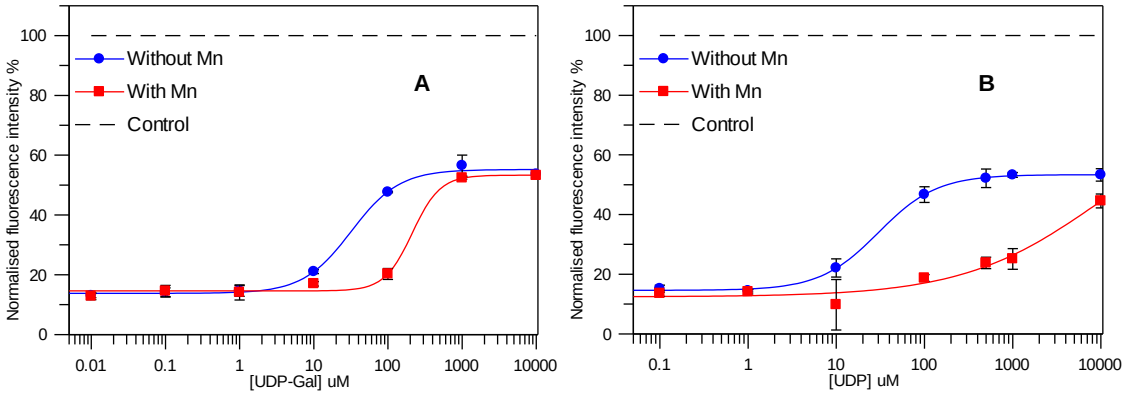


Figure 2.11. Effect of manganese ion (10 mM) in the ligand displacement assay with UDP-Gal (A) and UDP (B)

The manganese (II) ion activates various metal-dependent GTs[5, 26] and several GT biochemical assays are developed in the presence of excess of Mn^{2+} (mM range).[74, 82, 84] The maximum catalytic activity of GTs is reached faster in the excess of Mn^{2+} and thus biochemical assays can be run in lower enzyme concentrations. However, the sensitivity of the LDA was lost in the presence of 10 mM Mn^{2+} because the fluorophore was not displaced at low concentrations of UDP and UDP-Gal. The aim was to develop a sensitive assay to identify inhibitors. Based on the obtained results, if Mn^{2+} is added to the assay buffer then higher concentration of binder/inhibitor is required to displace the fluorophore and this may increase a possibility of false negative hits in inhibitor screening. Furthermore, if Mn^{2+} is not added, the reaction conditions are closer to natural human tissue Mn^{2+} concentration (µM range).[148] Because the sensitivity of the LDA was lost in the presence of Mn^{2+} all subsequent LDA experiments were performed without Mn^{2+} addition to the assay buffer.

2.2.4. Non-specific binding and influence of surfactant

All the binding affinity experiments indicated that certain non-specific binding of the fluorophore was occurring because the fluorescence did not quench linearly with increasing concentration of LgtC. The fluorescence quenched in two steps at low LgtC concentrations. With increasing concentration of LgtC (Figure 2.12, blue curve): the fluorescence initially was quenched sharply, (step 1), then levelled off (2) and quenched again, reaching plateau (3).

The fluorophore does not exert any specific binding to bovine serum albumin (BSA) and therefore, BSA was chosen for the investigation of non-specific binding. With increasing concentration of BSA the fluorescence initially quenched sharply and then completely levelled off at the same fluorescence intensity as the first plateau (step 2) was observed with LgtC (Figure 2.12). These results indicated that non-specific binding is not protein dependent and non-specific binding can possibly be explained with the characteristics of the fluorophore.

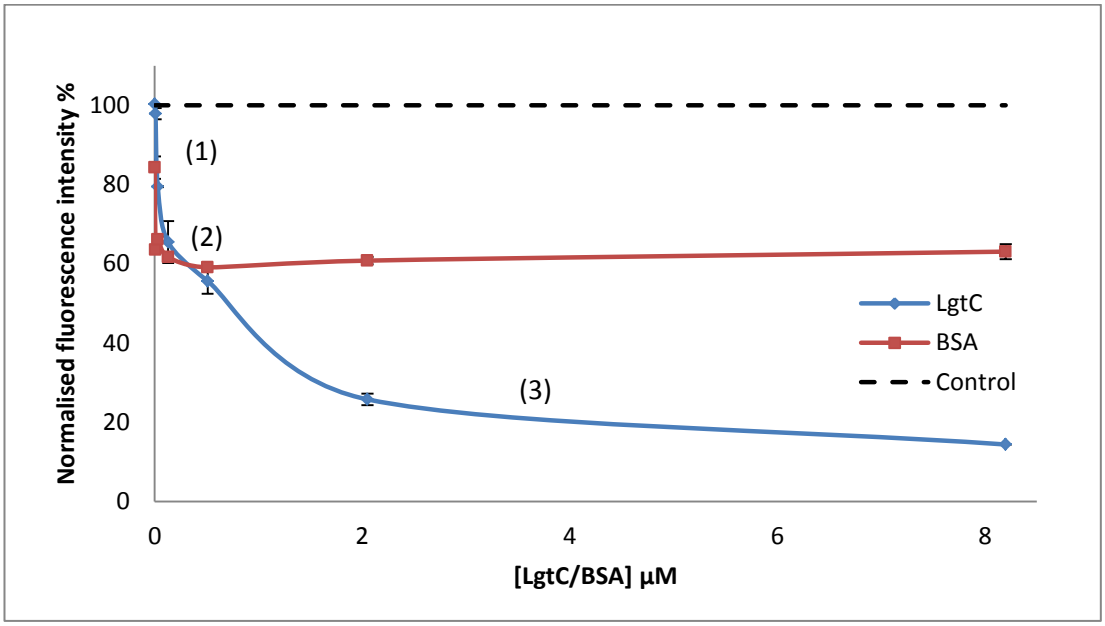


Figure 2.12. Investigation of non-specific binding with BSA as a control. Binding affinity of the fluorophore with increasing concentration of LgtC or BSA. Steps of quenching marked as 1-3 on the blue curve

A possible explanation for the "two-step" fluorescence quenching is that the fluorophore is able to aggregate and the fluorescence quenching may be due to the aggregate's non-specific binding to LgtC. The second quenching step (2) is possibly due to the actual fluorophore

binding to the active site of LgtC causing a decrease in fluorescence. These binding modes are likely occurring at the same time causing the "two-step" binding curve with increasing concentration of LgtC.

Triton X-100 (TX-100, polyoxyethylene octyl phenyl ether) is a non-ionic surfactant that can be used to protect proteins from binding to plastic surfaces (*i.e.* pipette tips or microplates) or dissolve molecular aggregators.[149] A critical micelle concentration (CMC) of TX-100 is 0.2 mM (approximately 0.02 % v/v).[150] The assay capability to tolerate TX-100 and how fluorophore binding changes in various concentrations of TX-100 (below CMC) was investigated.

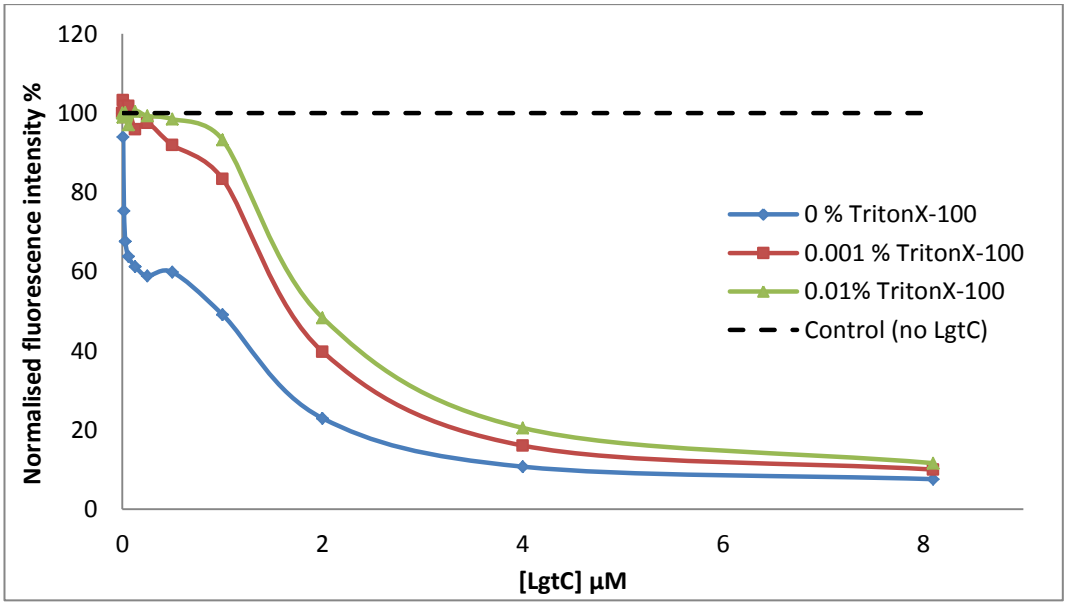
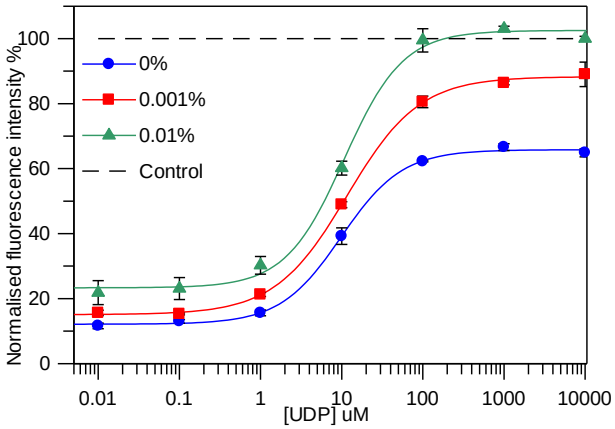


Figure 2.13. Fluorophore binding to LgtC in the presence of various TX-100 concentrations

The shape of the binding curves changed significantly in the presence of TX-100 (Figure 2.13). Now fluorescence initially remained stable, followed by fluorescence quenching and finally reaching a plateau. These results suggested that the first sharp quenching was due to non-specific binding of the fluorophore. This was then further investigated in the ligand displacement assay where the fluorescence was restored merely to sixty percent by addition of UDP. The sixty percent level was also observed from the binding affinity experiments (Figure 2.13, blue curve), where the first plateau was reached with increasing concentration of LgtC and in the absence of TX-100. Therefore, in order to investigate the influence of the surfactant the LDA was performed in various TX-100 concentrations.



[TX-100]-% (v/v)	IC_{50} (UDP), µM
0	10 ± 1
0.001	12 ± 1
0.01	11 ± 1

Figure 2.14. Influence on TX-100 at 0, 0.001 and 0.01% for fluorophore displacement with UDP and IC_{50} of UDP, respectively

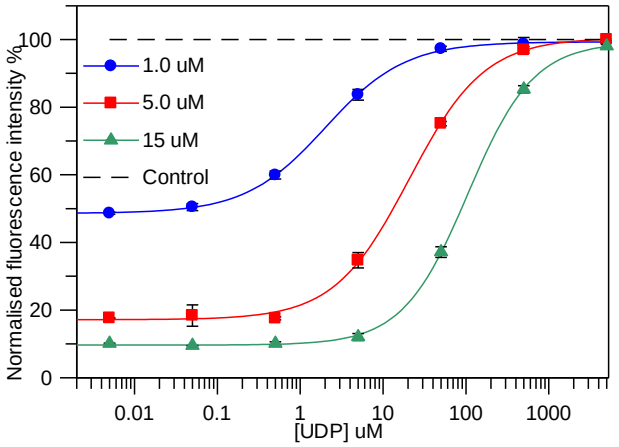
The results showed that the IC_{50} of UDP values were not affected by the TX-100 and as illustrated in Figure 2.14, the fluorescence was restored to 60 % in the absence of TX-100, 80 % at 0.001 % of TX-100 and completely restored at 0.01 % TX-100. These findings suggested that non-specific binding of the fluorophore is prevented in the presence of TX-100. The results also suggested that the fluorophore binds specifically to LgtC because the fluorescence can be completely restored with natural substrate which binds to the donor binding site. Thus, it was assumed that the fluorophore forms small particles which quench fluorescence due to unspecific binding. Therefore, fluorescence is impossible to restore completely with a natural donor which only displaces specifically bound fluorophore molecules. If the fluorophore forms small particles which TX-100 is able to dissolve then the completely restored fluorescence is observed.

These results were important, because the assay window (the fluorescence difference between negative and positive controls) increased in the presence of TX-100. Therefore, 0.01 % of TX-100 was added to the assay buffer in subsequent experiments. Furthermore, TX-100 and its capability to prevent compound aggregation are introduced in more detail in Chapter 4 where it will be used in compound library screening and inhibition characterisation studies.

2.2.5. Influence of LgtC concentration

An optimum concentration of LgtC was required for the LDA. LgtC concentration was chosen from the binding curve (normalised fluorescence intensity versus LgtC concentration) area where a maximum amount of fluorophore is bound to LgtC, approximately at 15 % normalised fluorescence intensity (Figure 2.13). At this LgtC concentration the maximum variation in fluorescence intensity can be observed in the fluorophore displacement experiments. It was assumed that a high LgtC concentration would increase and a low concentration would decrease IC_{50} values due to the availability of the number of donor binding sites. The LDA was performed in various LgtC concentrations in order to investigate how significantly IC_{50} changes.

As assumed, the IC_{50} increased significantly with increasing concentration of LgtC (Figure 2.15). Three fold increase in LgtC concentration raised IC_{50} from 21 µM to 106 µM. Most of the fluorophore is bound at concentrations between 5 and 8 µM of LgtC and fluorescence remains at the same level as observed in Figure 2.13. However, excess of LgtC concentration should be avoided because more binder would be required to displace the fluorophore and the assay would lose sensitivity.



[LgtC] µM	IC_{50}, µM
1.0	2.0
5.0	21
15	106

Figure 2.15. Fluorophore displacement with UDP at 1.0, 5.0 and 15 µM of LgtC, and IC_{50} of UDP respectively. The LDA performed without the addition of Mn^{2+} and in the presence of TX-100 (0.01 %)

The optimum LgtC concentration for subsequent LDA experiments was suggested to be optimised individually for each LgtC batch because the IC_{50} values are significantly dependent on the LgtC concentration. The concentration optimisation can be performed by the binding affinity assay with increasing concentration of LgtC and in fixed concentration of the fluorophore.

2.2.6. Summary of ligand displacement assay optimisation and determination of assay reproducibility

The LDA was developed for the discovery of novel inhibitors against GTs. LgtC was used for the assay optimisation and most of the assay parameters were changed from the original assay set up,[102] including reaction volume, buffer, concentration of DMSO, the addition of Mn^{2+} was omitted and TX-100 was introduced (Table 2.3). Overall, during the optimisation the LDA proved to be a very robust technique that is not affected by the assay conditions, by pH of the assay buffer or by the concentration of DMSO.

Table 2.3. Comparison of the "original" and the optimised LDA parameters

Parameter	"original"	"optimised"
Reaction volume	200 µL	100 µL
Buffer	50 mM Tris/HCl, pH 7.0	13 mM HEPES, pH 7.0
[Fluorophore][a]	200 nM	200 nM
Incubation time	10 minutes	20 minutes
Incubation temperature	30 °C	30 °C
[DMSO]	5 %[b, c]	10 %[b]
[$MnCl_2$]	10 mM	-
[LgtC]	unknown	Variable
[TX-100]	-	0.01 %[b]

[a]5FTUDP-Gal, [b]%-v/v, [c]only in compound screening experiments

The LDA (IC_{50} values) is sensitive to Mn^{2+} concentration and LgtC concentration changes. IC_{50} values increased significantly and the assay sensitivity was poorer when 10 mM $MnCl_2$ was added to the reaction buffer (Chapter 2.2.3). Therefore, addition of $MnCl_2$ was omitted in order to keep the assay more sensitive for compound screening. Variations in IC_{50} values from 10 µM to 44 µM with UDP were observed between LgtC batch changes (Table 2.4). LgtC is prone to precipitate after purification, therefore, the determined protein concentrations may have a slight variation. The optimal concentration of LgtC was determined prior to the IC_{50} experiments, by recording fluorescence intensity in various LgtC concentrations. The optimum concentration of LgtC was found when normalised fluorescence intensity quenched approximately to 15 %. Variable IC_{50} values hampered the comparison of the results, therefore, if results were compared then it was important that the same batch and concentration of LgtC was used.

Table 2.4. Concentrations of various LgtC batches and determined IC_{50} values with the optimised LDA

LgtC batch	[LgtC] stock	[LgtC] in assay	IC_{50} (UDP)
1	270 µM	13 µM	31 µM
2	135 µM	10 µM	44 µM
3	321 µM	4 µM	10 µM
4	225 µM	5 µM	21 µM

It is important to verify that the assay results are reproducible and the variability of the results of the assay is acceptably low. Assay parameters, including signal-to-background (S/B) ratio and Z' factor determine the assay robustness and reproducibility during normal assay conditions.[151] S/B ratio is the mean maximal signal and the mean minimum signal, describes the dynamic range of the assay without any data variation information and the Z' factor describes the available signal window which is the total separation between negative and positive control minus the error associated with each control. The Z' factor defined as the Equation 1, where std_a is the standard deviation of positive control (the highest obtained FI), std_b is standard deviation of a negative control (the lowest obtained FI), a is the mean of the positive control and b is the mean of the negative control.[151]

Equation 1. $$\mathbf{Z'} = \mathbf{1} - \left(\frac{(\mathbf{3std_a + 3std_b})}{(\mathbf{a - b})}\right)$$

In order to validate the LDA the reproducibility was examined by determining the S/B ratio and Z' factor with the optimised LDA conditions (Table 2.3). Two 96 well microplates were prepared on different days in the absence and in the presence of TX-100. Each microplate contained 32 replicate negative controls (fluorophore + LgtC), replicate positive controls (fluorophore + LgtC + 5 mM UDP) and control wells at the highest fluorescence (fluorophore).

Table 2.5. Z' factor and S/B ratio of the optimised LDA

Surfactant	Z' factor	S/B
0 % TX-100	0.54	5.9
0.01 % TX-100	0.73	7.7

An acceptable Z' factor is greater than 0.5 for HTS[151] and the assay reached this criteria in the absence and in the presence of TX-100, Z' factors 0.54 and 0.73 respectively (Table 2.5). TX-100 has a capability to lower the surface tension of aqueous solutions and this feature possibly had a positive effect on reproducibility because the Z' factor increased from 0.54 to 0.73 in the presence of TX-100. Additionally, as presented in the chapter 2.2.4, in the presence TX-100 the fluorescence intensity recovered completely which resulted in an increase to the S/B ratio from 5.9 to 7.7. High Z'- value (>0.5) and S/B ratio (>5) demonstrated that LDA was repeatable and robust for future HTS of compound libraries.

2.3. Stability of LgtC

Enzymes can lose their activity during storage due to conformational changes. The change in activity can be observed through activity assays where the substrate concentrations are constant, so the explanatory factor of the decrease is a reduction in the activity of enzyme. The stability and change in function of LgtC was investigated with four different methods after two and five months from the expression. Four different methods were chosen to ensure that the change in LgtC function is not caused by the standard deviation of the assay. LgtC was expressed and purified in house as previously reported.[16] Detailed expression and purification of LgtC is described in experimental chapter 2.7.1.

- Binding of the fluorophore was determined with the binding affinity assay
- Fluorophore displacement was determined by the LDA (IC_{50} of UDP-Gal)
- Specific activity was determined by NADH coupled assay directly after expression and after 5 months (see chapter 2.7.2)
- Binding affinity of UDP-Gal (K_m) and the catalyst rate constant (K_{cat}) were determined by phosphatase coupled assay (see chapter 3.4)

The results suggested that the fluorophore binds to LgtC with lower affinity after 5 months from expression than after 2 months. A change in binding curves was observed particularly between 1-5 µM of LgtC (Figure 2.16 A). Although the same concentration and batch of LgtC was utilised in the LDA, less UDP-Gal was required to displace the fluorophore (Figure 2.16 B) resulting decrease in the IC_{50} value from 32 µM to 13 µM. The results are comparable to the results in chapter 2.2.5, where various LgtC concentrations were examined and decreased IC_{50} values were observed at lower LgtC concentration.

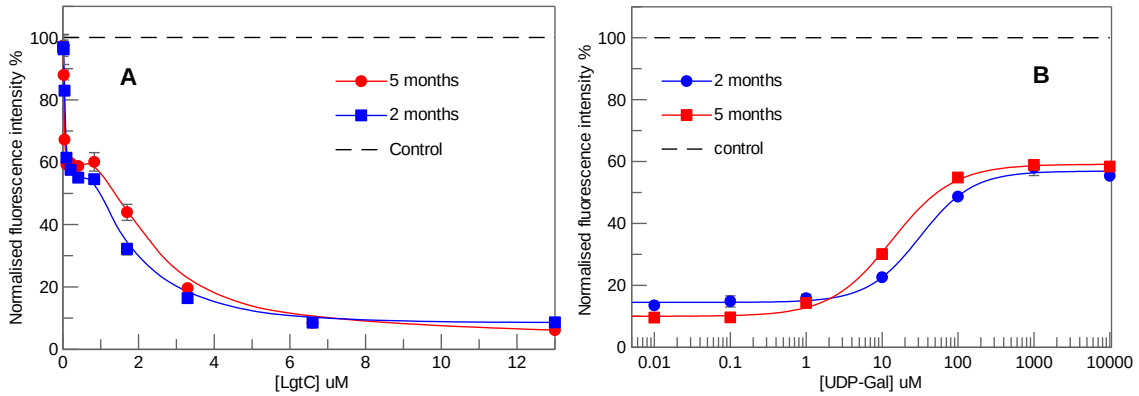


Figure 2.16. Change in binding affinity (A) and the fluorophore displacement (B) after 2 and 5 months expression of LgtC. IC_{50} of UDP-Gal decreased from 32 µM to 12 µM over the time

The investigations performed by two different biochemical assays indicated that catalytic activity of LgtC decreases over the time. The specific activity of LgtC dropped from 4.2 U/mL to 3.7 U/mL (12 %) in five months. In addition, K_m and K_{cat} values changed significantly (Table 2.6). K_m increased from 27 µM to 81 µM and LgtC capability to turn substrate molecules to product reduced from 0.12 to 0.05 per second.

Table 2.6. Change in specific activity, K_m, K_{cat} and IC_{50} of UDP-Gal determined with LgtC after 2 and 5 months after the expression

Storage time	Specific activity of the stock[a]	K_m [b]	K_{cat} [b]	IC_{50} (UDP-Gal)[c]
2 months	4.2 U/ mL	27 µM	0.12 s^{-1}	32 µM
5 months	3.7 U/ mL	81 µM	0.05 s^{-1}	12 µM

[a]NADH coupled assay (Chapter 2.7.2), [b]phosphatase coupled assay (Chapter 3.4), [c]optimised LDA as described in Table 2.3 except in the absence of TX-100

All the results indicated that function and catalytic activity of LgtC is reduced over the examined time, possibly by denaturation. As all three assays indicated lower binding affinity or lower activity actions of LgtC over time, the observed changes were most likely due to changes in LgtC, not the standard deviation of the assay used. The results suggested that a conformational shift of LgtC changes the binding affinity of the substrates over the time. LgtC preserves its function for less than five months, therefore, regular binding affinity and activity controls were suggested. If the IC_{50} values were compared then the binding affinity of the fluorophore experiment should be performed prior to the LDA to choose the optimal LgtC concentration. If the impact of an inhibitor to K_m is investigated then the K_m should be

determined in the absence of inhibitor on the same day as a control. These control experiments are important as the use of aged LgtC may suggest inhibition (false positive results), although the change in activity can be caused by an inactive or less active LgtC.

2.4. Assay adaptation to 384 well microplates

HTS approaches are commonly used in drug discovery and 384 well microplate formats are widely used.[152, 153] Compared to 96 well formats, the 384 well formats allow the determination of numerous compounds simultaneously and a much smaller amount of valuable materials are used. These factors contribute to making large compound library screening more practically viable. All the assay optimisation and control experiments were performed in 96 well microplates. Therefore, the aim was to adapt the 96 well format to 384 well microplates. To start with, the reaction volume of two different 384 well microplates were evaluated and compared. The previously optimised assay parameters were then directly adapted to 384 micro well format.

2.4.1. Linearity of fluorophore and reaction volume

Non-coated 384 well microplates from two different suppliers, Nunc and Greiner, were compared. Nunc 384 well microplates were made from polystyrene the same as the Nunc 96 well microplates and Greiner microplates were made from polypropylene. The functionality of the microplates was evaluated by determining a concentration range of the fluorophore (5FTUDP-Gal) in which the concentration and the measurement signal has a linear dependence: linear regression (r^2). The linear regression of the fluorophore was determined in various reaction volumes with Nunc and Greiner microplates to attain the optimum reaction volume and microplate for the LDA experiments.

Table 2.7. Linearity of fluorophore in range of 12.5 to 1000 nM (10 measurement points) in various reaction volumes performed in the assay buffer (13 mM HEPES pH 7.0)

Volume in well, µL	r^2 (Nunc)	r^2 (Greiner)
10	0.998	1.000[a]
20	0.996	0.999
30	0.999	0.999
40	0.999	0.998
50	0.996	0.999

[a]Volume 15 µL in the well (the lowest recommended volume by the manufacturer)

The linearity of the fluorophore was identical with both 384 well microplates (Table 2.7). Greiner micro plates were matte finished, practically easier to work with and less expensive than Nunc microplates. Therefore, Greiner microplates were chosen for subsequent experiments. Reaction volume did not impact the fluorescence of the fluorophore linearity and because the lowest possible reaction volume was desirable, 20 µL was selected.

2.4.2. Fluorescence based ligand displacement assay

Previously optimised LDA parameters (Chapter 2.2.6) were directly adapted to 384 well microplate format. The binding affinity of the fluorophore was identical when the binding curves obtained with the 96 and 384 well microplates were compared (Figure 2.17 A). Consequently, the same concentration and the batch of LgtC were used for the ligand displacement experiment (Figure 2.17 B) in 96 and 384 well microplates.

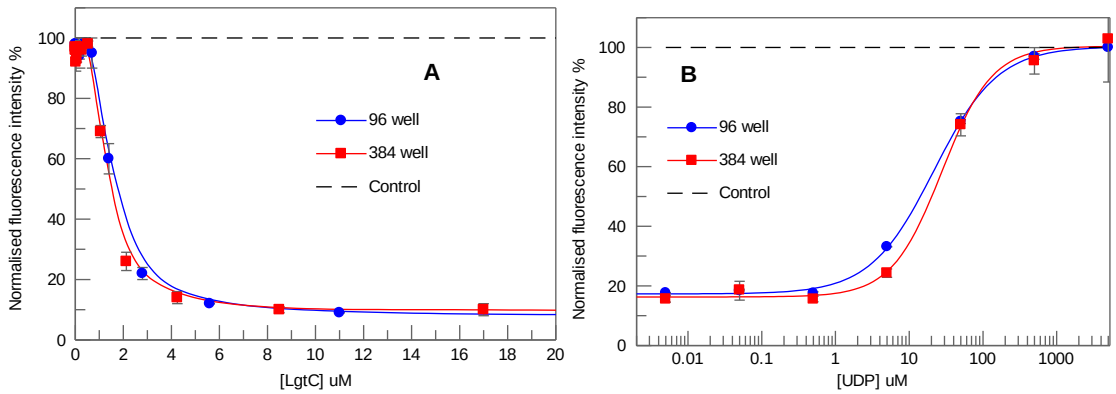


Figure 2.17. Comparison of 96 and 384 well microplates. Binding of fluorophore increasing concentration of LgtC (A) and ligand displacement with UDP (B) determined by the optimised assay parameters (Chapter 2.2.6). IC_{50} of UDP 28 ± 4 µM and 22 ± 1 µM in 384 and 96 well microplates, respectively.

The results indicated that the studied 384 well microplate functioned in a similar fashion to the 96 well microplate and further optimisation was not required. The 384 well microplate is practically useful when only a small amount of material (*i.e.* protein) is available or equally importantly, for compound library screening.

2.4.3. Ligand displacement with other sugar nucleotides and nucleotides

The suitability of the LDA for compound screening was investigated with various sugar nucleotides and nucleotides that have different binding affinity to LgtC. UDP-Gal is a natural donor substrate to LgtC[16] and it displaces the fluorophore in competitive fashion. Also UDP, UMP and uridine are assumed to bind specifically to the donor binding site of LgtC and behave as competitive inhibitors, although with lower affinity than UDP-Gal.

The ligand displacement experiment was performed in fixed concentrations of all assay components except the binders, UDP-Gal, UDP, UMP and uridine on the 384 well microplate format and with the optimised assay conditions (Chapter 2.2.6). The IC_{50} values were determined for each binders and then the obtained IC_{50} values were compared with the published results which were performed with the preliminary assay conditions.[104]

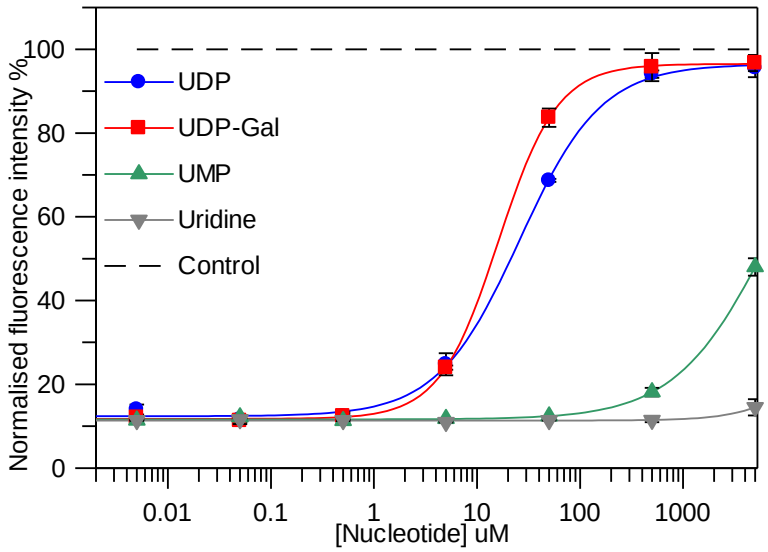


Figure 2.18. Fluorophore displacement with UDP-Gal, UDP, UMP and uridine determined with the optimised LDA (DMSO (10 %), TX-100 (0.01 %) in 13 mM HEPES (pH 7.0) buffer)

Results indicated that UDP-Gal binds to LgtC with the highest affinity then UDP, UMP and finally uridine as expected (Figure 2.18). The same trend was observed with the published results (Table 2.8). A variation between the determined and the published IC_{50} values was observed. The main difference between the assays was the concentration of $MnCl_2$ (published results[104]), and the LDA was performed in the presence of DMSO and TX-100. As previously discussed in chapter 2.2.6, the IC_{50} values are mostly affected by the concentration of Mn^{2+} and LgtC, however, not the concentration of DMSO or TX-100 (chapters 2.2.1 and 2.2.4, respectively). Slightly different IC_{50} values can be explained with LgtC concentration used in the assay. The LDA assay was performed at 8 µM of LgtC, however, the concentration was unknown in the reported assay. The most significant difference between the results is the lower IC_{50} values in the presence of Mn^{2+}. As demonstrated in the chapter 2.2.3, IC_{50} of UDP-Gal increased from 32 µM to 213 µM in the presence of 10 mM Mn^{2+}. The possible explanation for different IC_{50} values in the presence of Mn^2 (26 µM vs. 213 µM for UDP-Gal) is the actual concentration of Mn^{2+} and it could be that the concentration of Mn^{2+} was lower than 10 mM in the assay of the published results.

Table 2.8. IC_{50} values for various nucleotides determined the optimised LDA and the published results of the same experiment (both assay conditions described in detail in Chapter 2.2.6)

Nucleotide	IC_{50}, µM[a]	IC_{50}, µM[b]
UDP-Gal	16 ± 1	26 ± 8
UDP	26 ± 2	83 ± 49
UMP	< 1000	293 ± 91
uridine	< 1000	< 1000

[a]Optimised assay (10 % DMSO, 0.01 % TX-100 in 13 mM HEPES pH 7.0 buffer, 384 well microplate), [b]Published results (10 mM $MnCl_2$ in 50 mM Tris/ HCl pH 7.0 buffer, 96 well microplate)[104]

Overall, the results showed that the compound screening assay can be performed in 384 well microplates and how low affinity binders would behave in the LDA. In addition, there is a difference in the IC_{50} values obtained with the optimised assay and the preliminary assay set up[102, 104] which is a result of the different assay conditions used.

2.5 Investigation of other glycosyltransferases

The binding affinity experiments demonstrated that the fluorophore binds to LgtC. The specific binding was demonstrated with the natural donor substrate on the ligand displacement experiments. Previously, it has been shown that quenching is not limited to LgtC. The fluorophore specifically bound also to other GTs resulting fluorescence quenching: α1,3-GalT from *B. Taurus* (α1,3-GalT) and human blood group B enzyme (GTB) and AA(Gly)B (GTB mutant).[104] The studied GTs belongs to the GT family 6, except LgtC, and all possess a retaining stereochemistry and a GT-A fold. Therefore, the possibility to expand the use of the fluorescent donor analogues (5FTUDP-Gal and 5FTUDP-Glc) and the LDA for other GTs was investigated. GTs with different characteristics were chosen, including different GT families, GT folds and stereochemistry (Table 2.9). The availability of UDP-Glc analogue, (5-(5-formyl-2-thiophene)-UDP-Glc, (5FTUDP-Glc, Figure 2.19) allowed the investigation of two bacterial GlcTs, NGT and TcdB. Previous studies showed that the fluorophore (5FTUDP-Gal) behaved as a donor substrate for mammalian GT, β1,4-GalT, demonstrating the binding ability of the fluorophore to this GT.[102] Therefore, β1,4-GalT was included to the binding affinity experiments as a control.

Table 2.9. Properties of studied GTs

GT	Transferred sugar	GT family	Organism	Stereochemistry	Fold
LgtC	Gal	GT8	*N. meningitidis*	Retaining	GT-A
TcdB	Glc	GT44	*C. difficile*	Retaining	GT-A
NGT	Glc and Gal[a]	GT41	*A. pleuropneumoniae*	Inverting	GT-B
β1.4-GalT	Gal	GT7	*B. taurus*	Inverting	GT-A

[a]NGT transfers galactose with lower affinity

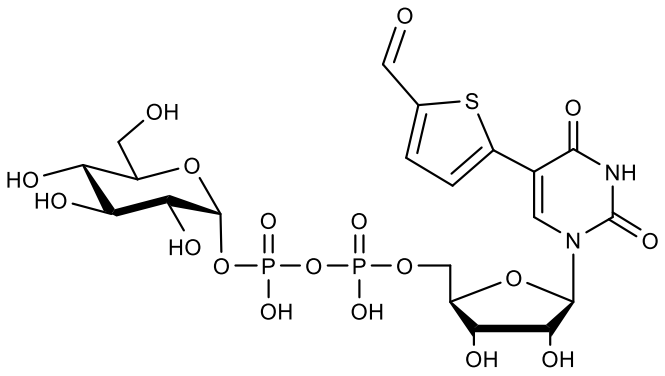


Figure 2.19. Structure of fluorescent UDP-Glc analogue (5-(5-formyl-2-thiophene)-UDP-Glc, 5FTUDP-Glc) used studying NGT and TcdB

The binding affinity of the fluorophores (5FTUDP-Gal/ Glc) to β1,4-Gal, NGT and TcdB was investigated and compared with the binding curve of LgtC. The experiments were performed using similar conditions as for LgtC. Buffer and pH were specifically adapted for each enzyme. For example, TcdB was found to be active in buffer containing 25 mM Tris, 150 mM NaCl, 150 mM K_2SO_4, 5 mM $MnCl_2$ and 5 mM $MgCl_2$ at pH 7.5[85], therefore, this particular buffer was utilised in the binding affinity experiments.

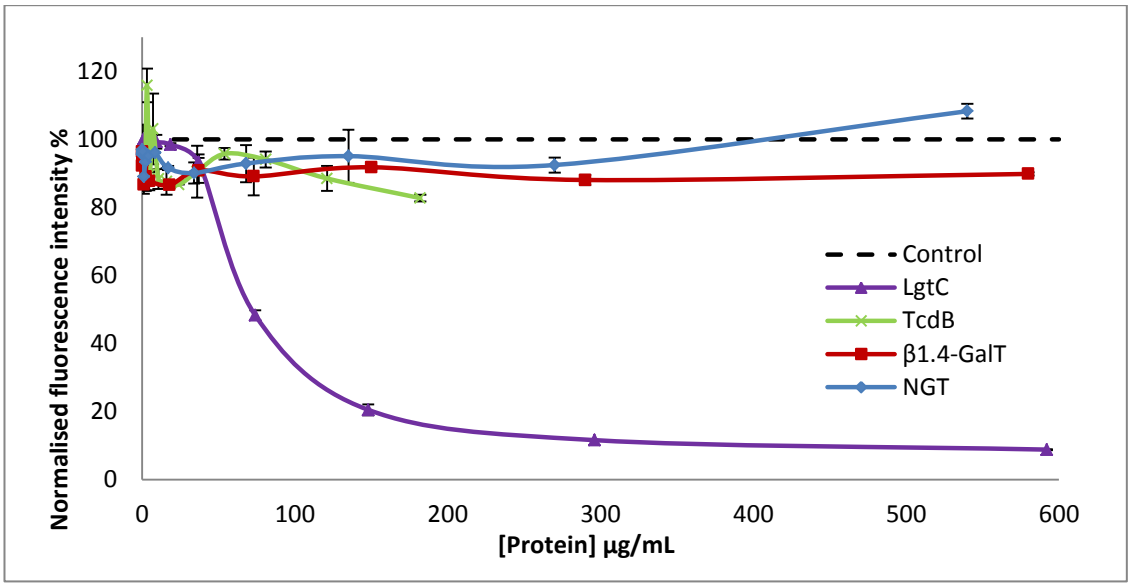


Figure 2.20. Fluorophore binding monitored with increasing concentrations of LgtC, TcdB, β1.4-GalT and NGT. The experiments performed in the presence of TX-100 and DMSO

The results were unexpected because fluorescence quenching was not observed with other GTs: TcdB, β1.1-GalT or NGT (Figure 2.20). The previous enzyme kinetic studies strongly suggested that the 5FTUDP-Gal binds to the β1,4GalT[102] however, the fluorescence remained high. Possible explanations for the observed results are that binding of the fluorophore does not occur, binding occurs but fluorescence intensity remains high or protein concentration is too low.

Hypothetically, the fluorescence quenching could be explained by the coverage of the fluorophore. All examined GTs contain flexible loops that are located near the active site and during donor sugar binding the loops are suggested to become active participating in donor binding and product release.[16, 120, 134, 141] In case of LgtC and TcdB the sugar moiety of the donor is bound to the bottom of the active site (Figure 2.21 A and D), under the phosphate groups.[16, 154] In the case of LgtC, the loop area covers the sugar donor within the enzyme,[16, 108] whereas with TcdB, the rest of the donor sugar (UDP) is exposed to the solvent.[134] Similar

movements have been observed in NGT and β1,4-GalT.[120, 141] UDP is almost completely buried in the cleft of NGT (Figure 2.21 D).[141] This could promote fluorescence quenching however, the 5-substituted formyl-2-thiophene may prevent subtle conformational changes that further precludes 5FTUDP-Glc binding and fluorescence remains high. The active site of β1,4-GalT is large (Figure 2.21 B) and during the sugar donor binding the acceptor binding site is created, leaving the sugar donor exposed to the solvent.

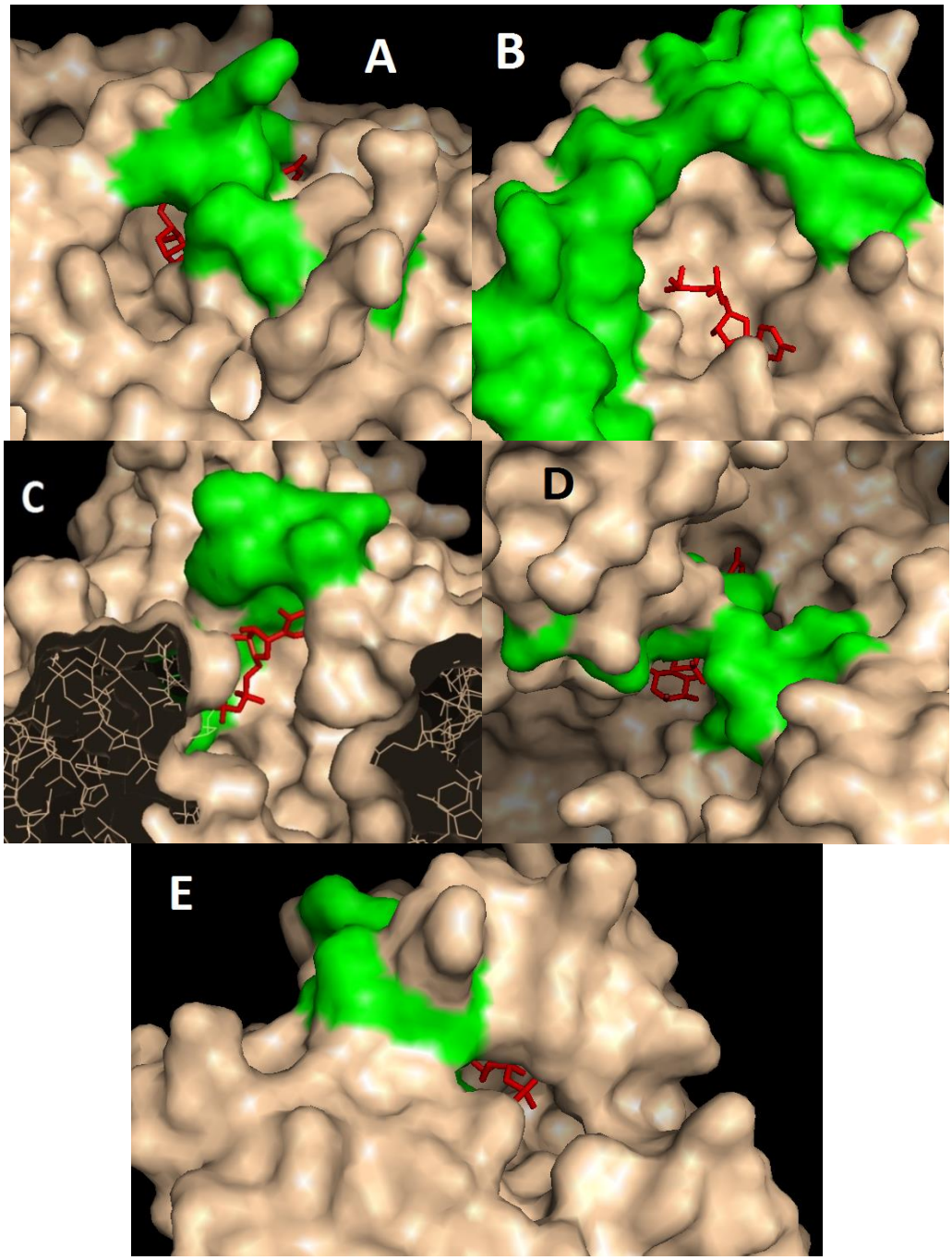


Figure 2.21. Molecular surface presentation of the active sites of A) LgtC (PDB: 1G9R), B) β1,4-GalT (PDB: 1FR8), C) NGT (PDB: 3Q3H), D) TcdB (PDB: 2BVL), and E) α1,3-GalT (PDB: 1K4V) in the presence of donor substrates. Donor substrates (red) are shown in stick form and loop regions are shown in green

The amino acids involved in binding of the sugar donor and with uracil are diverse. It is not clear which amino acids could interact with the 5-formyl-2-thiophene and cause the fluorescence quenching. Recently reported results demonstrated that fluorescence intensity was quenched when 5FTUDP-Gal was incubated with bovine α1,3-GalT.[104] The amino acids interacting with the uracil portion of α1,3-GalT and LgtC are different, and it cannot be concluded which amino acids could promote fluorescence quenching. For example, valine 136 interact with O2 and N3, water molecule interact with O4 of the uracil of α1,3-GalT[155] and aspartic acid 8 interact with O2 and N3, asparagine 10 interact with O4 of the uracil of LgtC.[16] However, the donor substrate is buried within both GTs, LgtC and α1,3-GalT (Figure 2.21, A and E, respectively) that could quench the fluorescence signal. Mutagenesis studies with α1,3-GalT suggested that binding of sugar donor induces conformational changes in both the ligand and in the enzyme. Studies indicated that UDP-Gal can bind to α1,3-GalT in a distorted conformation[156] similarly than UDP-2FGal to LgtC wherein pyranose ring is bent under the phosphates and is nearly parallel to the plane of the pyrophosphate.[16] Hypothetically, this conformational change could occur in the fluorophore upon binding to LgtC however, not with β1,4-GalT. One possibility is that 5FTUDP-Gal is a so called "molecular rotor". Molecular rotors are fluorescent molecules that undergo intramolecular twisting motion upon photon absorption.[157] Non-planar (twisted) state has lower excited energy and higher ground stage energy, therefore, the corresponding lower relaxation energy. Dependent on the chemical structure, relaxation from the twisted state is associated either with longer emission wavelength or without photon emission (non-fluorescent).[157] Therefore, possibly the planarity of the thiophene-uracil in bound 5FTUDP-Gal is disturbed due to binding which may lead to twisted intramolecular charge transfer and quenching in fluorescence.

Alternatively, during the fluorophore binding to NGT and TcdB, the fluorophore may prevent conformational changes of the enzymes and the suggested loops remain inactive or the structure of the fluorophore is unfavourable and binding to the donor binding site does not occur. Alternatively, higher concentrations of NGT and TcdB may be required to obtain fluorescence quenching. However, this is unlikely if the fluorophore binds to other GTs similarly than to LgtC because, comparing the concentrations used in the binding affinity assay, at 150 μM/mL of LgtC fluorescence was significantly quenched however, quenching was not observed with other GTs.

The results suggests that binding of the fluorophore can occur although the fluorescence quenching is not observed because the previous studies strongly indicated that the 5FTUDP-Gal served as substrate of β1,4-GalT.[102] Hypothetically, the fluorescence intensity is quenched when the whole fluorophore is buried within the enzyme or alternatively, the fluorophore molecule is twisted during favourable binding to active pocket of the enzyme. Currently, the fluorophore, 5FTUDP-Gal, can only be used in the investigation of LgtC, α1,3-GalT and GTB.

2.6 Summary and conclusions

A robust and operationally simple assay for compound screening against glycosyltransferases was successfully optimised from a previously described fluorescence based ligand displacement set up[104] by using LgtC as a model enzyme. The majority of the optimised assay conditions were changed from the preliminary method (Table 2.3).

The most significant assay conditions that effect the assay results were $MnCl_2$ and enzyme concentrations. At high $MnCl_2$ concentration, a higher concentration of binder was required to displace the fluorophore (5FTUDP-Gal) and because the sensitivity of the assay reduced, $MnCl_2$ was omitted. In high enzyme concentration availability of donor binding sites increased and more fluorophore was bound which required higher concentrations of binders to displace the fluorophore. The optimum enzyme concentration was found when normalised fluorescence intensity was quenched approximately to 15 %. In addition, the LDA (IC_{50} values) was demonstrated to be sensitive to the concentration of LgtC (chapter 2.2.6). Therefore, if the analysis of two sets of results were compared, then the same batch and concentration of LgtC was used in each experiment. The optimum concentration of LgtC is also required for the identification of potent inhibitors because too high concentration of enzyme could cause false negatives. The LDA uses a high concentration of LgtC compared to biochemical assays, where the required concentration is approximately 100 fold lower. The difference can possibly be explained by the assay mode. In the LDA, only binding of the fluorophore is monitored whereas in the biochemical assay, the formation of reaction product is monitored. In the biochemical assay one LgtC molecule produces numerous product molecules whereas in the LDA, the balance between the fluorophore and LgtC is

observed, therefore, a higher concentration of LgtC is required in order to obtain desired results.

The surfactant, TX-100 had a positive impact on the LDA because in the presence of TX-100 the characters of non-specific binding disappeared: the “two-step” binding curves and incomplete fluorescence recovery with specific binders. As fluorescence was completely restored in the presence of TX-100 the assay signal window improved (S/B ratio). The influence of TX-100 was also observed with increased Z’ factor that indicated improved reproducibility of the LDA. Hypothetically, the non-specific binding was caused by aggregation of the fluorophore which TX-100 dissolved.

The stability and storage time was unknown for LgtC. The change in specific activity, IC_{50} and K_m values was determined by NADH coupled assay (described in chapter 2.7.2), the LDA and the phosphatase coupled assay (described in chapter 3.4), respectively for the same LgtC batch 2 and 5 months after the expression. The results were important because the binding affinity of the fluorophore and the activity of LgtC changed over the examined time period. Therefore, a control that measures the function of LgtC was suggested to be included in the experiments.

Other GTs were examined in order to expand the use of fluorophores (5FTUDP-Gal/Glc) and investigate the applicability of the LDA. Fluorescence quenching was not observed with NGT, TcdB or β1,4-GalT. The binding mechanisms of the fluorophore and which amino acids are involved in fluorescence quenching are currently unknown. Previous experiments demonstrated that β1,4-GalT can use the fluorophore (5FTUDP-Gal) as a donor substrate,[102] suggesting that the binding of the fluorophore can occur without fluorescence quenching. Based on the coverage of the fluorophore within the catalytic pocket, the fluorescence intensity is quenched possibly when the whole fluorophore is buried within the enzyme. Alternatively, the planarity of the fluorophore is lost upon binding which results in fluorescence quenching. Currently, there is no clear trend for fluorescence quenching based on GT-family and the binding of the fluorophore is observed with retaining and inverting GTs using other methods. In practise, the usage of the fluorescent UDP-Gal /Glc analogues and the LDA is currently limited. GTs which can be investigated by the LDA are presented in Table 2.10.

Table 2.10. Summary of all GTs investigated with the binding affinity experiments in the current and in the previous study

GT	Transferred sugar	GT family	Stereo-chemistry	Fold	Fluorescence quenching?	Proved binding?
LgtC[a]	Gal	GT-8	Retaining	GT-A	yes	no
TcdB[a]	Glc	GT-44	Retaining	GT-A	no	no
NGT[a]	Glc	GT-41	Inverting	GT-B	no	no
β1.4-GalT[a, c]	Gal	GT-7	Inverting	GT-A	no	yes
α1,3-GalT[b, c]	Gal	GT-6	Retaining	GT-A	yes	yes
GTB[b]	Gal	GT-6	Retaining	GT-A	yes	no
AA(Gly)B GTB mutant[b]	Gal	GT-6	Retaining	GT-A	yes	no

[a]Current study, [b]previous study[104], [c]previous study[102]

Overall the LDA proved to be a very robust method during the assay optimisation. The determined IC_{50} values remained unchanged in the studied assay buffers and pHs, concentration of DMSO or incubation time. The Z' factor (>0.5) that is generally used parameter for checking the quality of screening assays, illustrated a high reproducibility of the LDA. The LDA is practical to perform because the assay format does not include acceptors and the enzymatic reactions are carried out in a low reaction volume (100 µL) in 96 micro well plates that reduce the consumption of valuable reagents. Additionally, the LDA was adapted to 384 micro well plates which further reduce the consumption of the assay reagents and make the assay more suitable for large compound library screening. The simple assay protocol contains only three pipetting steps, and together with 384 well microplates, the assay can be simply adapted to HTS. The optimised assay was developed for inhibitor screening and these experiments are described in chapter 4.

2.7 Experimental

2.7.1 Expression and purification of LgtC

The expression and purification process of LgtC followed a protocol that was obtained in collaboration with M. Palcic. The *Escherichia coli* (*E. coli*) cells containing LgtC plasmid were grown in tryptone broth (TB) + M9 salts + supplements (1M $MgCl_2$, 0.1M $CaCl_2$, 10mg/ml vitamin B1, 20% w/w glucose, 20% w/w casamino acids and 100 mg/ml ampicillin) liquid media in a shaker at 30 °C and 200 rpm for 16 hours. The main culture was prepared by diluting the starting culture with TB + M9 media to an OD_{600} of 0.1-0.2 to ensure that the cells are fresh. The main culture was then grown to an OD_{600} of 0.6-1.0 and the expression was induced with isopropylthio-β-galactoside (IPTG) at a final concentration of 0.5 mM. After incubation of the main culture at 30 °C for 16 hours, the cells were harvested by centrifugation at 4k rpm for 10 minutes at 4 °C and the cell pellet was re-suspended in disruption buffer (20mM MOPS, 500mM NaCl, 5mM imidazole, 1 protease inhibitor tablet (without EDTA), pH 7.0 adjusted with 1M NaOH). A sonicator probe at 50% duty cycle and at power setting 7 was used to disrupt the cells at 5 °C and the cell debris was removed by centrifuging at 20 rpm for 90 minutes at 4 °C. The supernatant was collected and filtered through 0.2 μm filters.

The supernatant solution was purified by IMAC on HisTap HP column with Ni^{2+}. The manual mode of GradiFrac was used combined with UV and conductivity detector. All buffers and the column were kept in ice during the purification. The column was equilibrated in loading buffer (composition of the buffers on Table 2.11) and non-specifically bound proteins were washed in washing buffer. The fractions were eluted with eluting buffer and collected in $MnCl_2$ solution to give final concentration of 5mM.

Table 2.11. The compositions of loading, washing and eluting buffers used in LgtC purification

Component	Loading	Washing	Eluting
MOPS	20 mM	20 mM	100 mM
NaCl	500 mM	500 mM	500 mM
Imidazole	5 mM	20 mM	500 mM
pH (adj. With NaOH)	7.5	7.5	no adjustment

The flow-through (fractions that were collected during loading the column), wash and protein fractions were analysed with SDS-PAGE. NuPAGE MOPS SDS buffer kit was utilised during the sample preparation. The samples were denaturated in NuPAGE LDS sample buffer in the presence of a reducing agent by heating the mixture to 70 °C for 10 minutes. 10 µL of the denaturated samples were loaded to the NuPAGE 4-12% Bis-Tris 10 well gels and the gels were placed into an electrophoresis chamber that was connected to PowerEase 500 Power Supply. The chamber was filled with MOPS running buffer and antioxidant solution was added to the middle chamber. The gels were run for 50 minutes at a voltage of 200 V and a current of 50 mA. After the run, the gels were stained with Simply Blue™ stain and analysed (Figure 2.22).

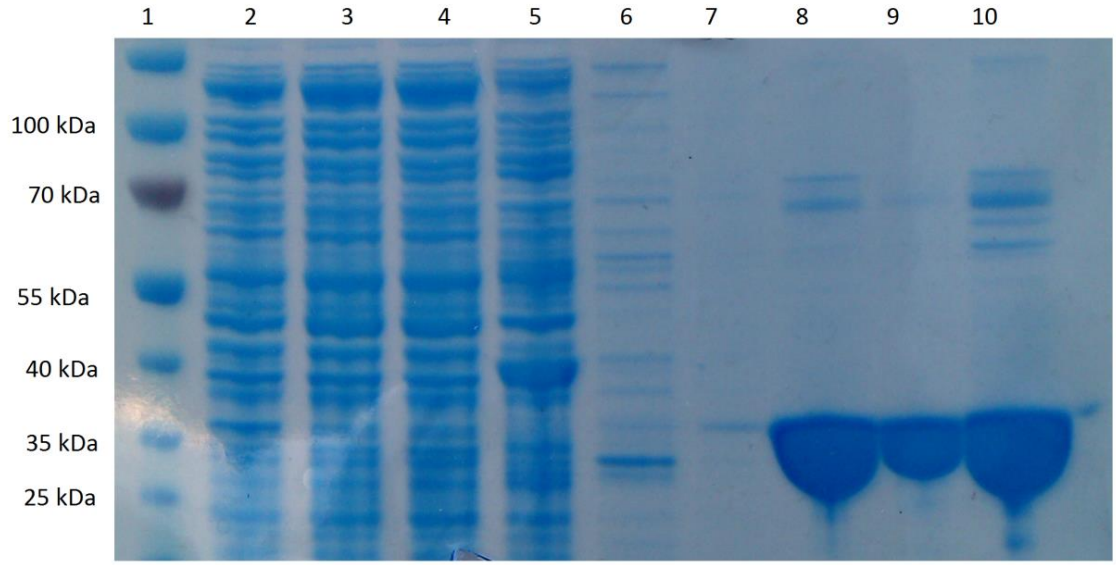


Figure 2.22. SDS-PAGE analysis of collected fractions. Lane 1 is a protein ladder. Lanes 2-3 are flow-through, lanes 4-6 are washes and 7-9 are collected LgtC fractions. Lane 7 is the first collected fraction (concentration not determined), lane 10 is LgtC from previous batch. Concentration of LgtC on lanes 8-10 are 11.84, 5.6 and 19.55 mg/mL respectively.

The fractions that contained a reasonable amount of pure LgtC were dialysed in 20 mM MOPS and 5 mM $MnCl_2$ buffer. After the dialysis LgtC was aliquoted and stored at -80°C.

2.7.2 Concentration and activity of LgtC

LgtC concentration of the stock solution was determined by a NanoDrop spectrophotometer. The activity of new LgtC batches were determined with a continuous assay where UDP is coupled to nicotinamide adenine dinucleotide (NADH) oxidation via pyruvate kinase (PK) and lactate hydrokinase (LDH).[84] The NADH coupled assay was optimised for determination of LgtC activity in house.

The activity assay was performed in clear 96 well microplates in total volume of 300 µL. All dilutions were made in 13 mM HEPES buffer at pH 7.0 including 50 mM KCl and 13 mM $MnCl_2$. A serial dilutions of concentrations of LgtC were activated in 5 mM DTT at 30 °C for 30 minutes. The reaction mixture contained PK (10 U/mL), LDH (15 U/mL), phosphoenolpyruvate (PEP) (700 µM), NADH (600 µM) and lactose (300 µM) (all final concentrations) was added to LgtC. The reaction was started by adding 300 µM of UDP-Gal to the reaction mixture and reduced absorbance at 340 nm was recorded. The assay contained positive control (UDP added instead of UDP-Gal), negative control (No UDP or UDP-Gal added) and blank (no NADH added). The rate of reaction in mAbs/min can be converted to mmol/min by using the standard curve of NADH concentration against absorbance. The unit activity (U/mL) and specific activity (U/mg) of the undiluted stock can be calculated by correlating with protein concentration.

2.7.3 General settings

All measurements were carried out in black NUNC F96 MicroWell polystyrene plates on a BMG labtech PolarStar microplate reader. The fluorescence intensity measurements were carried out at the excitation wavelength 350 ± 5 nm and an emission wavelength 430 ± 5 nm which were relevant to the fluorophores, 5FTUDP-Gal and 5FTUDP-Glc (λ_{ex} 351 nm, λ_{em} 434). 5FTUDP-Gal and 5FTUDP-Glc were synthesised and characterised in house by Dr Lauren Tedaldi. The flashes per well was set to 10 and gain was set to 90 % to the control (fluorophore only in assay buffer) well on the microplate. LgtC was expressed in house (see chapter 2.7.1) and molecular weight of 36000 Da[18] was used for calculating molar concentrations. All reagents used in this chapter are listed in App. 2.

50 mM Tris / HCl buffer and 13 mM HEPES buffers were prepared in ultrapure water and pH was adjusted with HCl or NaOH respectively. All buffers were filtered with nylon 0.2 µm membranes prior to use and stored in the refrigerator for maximum two weeks.

All assay components were prepared in 13 mM HEPES buffer at pH 7.0, unless otherwise stated. Fluorophore concentration is fixed (0.2 µM) in all experiments. Binder (*i.e.* UDP) concentration varied in the ligand displacement experiments whilst the enzyme concentration remained fixed. The binding affinity experiments were performed in variable enzyme concentration (Table 2.12). The mixture of fluorophore and TX-100 (+ DMSO in binding affinity experiments) in 13 mM HEPES buffer was added to the microplates first, then binder (only in the ligand displacement assay) and finally the enzyme. The final reaction volume was 100 µL. The samples were prepared at least in duplicate and incubated for 20 minutes at 30 °C before taking the final reading.

Table 2.12. Concentrations of the assay components in the binding affinity and ligand displacement assays

Component	[Stock]	[Final]	Volume of stock, µL
Binder[a]	Various [binder], x10 in 100 %[c] DMSO	Various [inhibitor] in 10 %[c] DMSO	10
Fluorophore	1 µM	0.2 µM	20
Enzyme[b]	X5	X1	20
HEPES buffer	13 mM	13 mM	40
TX-100	0.1 %[c]	0.01 %[c]	10

[a]In the binding affinity experiments binder is replaced with 100 % DMSO. [b]In the binding affinity assay enzyme concentration is variable whilst fixed in the ligand displacement experiments, [c]v/v-%

The general microplate layout in the binding affinity and ligand displacement experiments are presented in Figure 2.23 and Figure 2.24 respectively.

	1	2	3	4	5	6-12
A	M	M	M	M	M	...
B	M+E1	M+E1	M+E1	M+E8	M+E8	...
C	M+E2	M+E2	M+E2	M+E9	M+E9	...
D	M+E3	M+E3	M+E3	M+E10	M+E10	...
E	M+E4	M+E4	M+E4	M+E11	M+E11	...
F	M+E5	M+E5	M+E5	M+E12	M+E12	...
G	M+E6	M+E6	M+E6	M+E13	M+E13	...
H	M+E7	M+E7	M+E7	M+E14	M+E14	...

Figure 2.23. General microplate layout in the binding affinity experiments (96 well microplate). Enzyme (E) and mixture containing fluorophore, DMSO and TX-100 (M). All dilutions in 13 mM HEPES buffer (pH 7.0)

	1	2	3	4	5	6-12
A	M+E	M+E	M	M	…	…
B	M+E+I1	M+E+I1	M+I1	M+I1	…	…
C	M+E+I2	M+E+I2	M+I2	M+I2	…	…
D	M+E+I3	M+E+I3	M+I3	M+I3	…	…
E	M+E+I4	M+E+I4	M+I4	M+I4	…	…
F	M+E+I5	M+E+I5	M+I5	M+I5	…	…
G	M+E+I6	M+E+I6	M+I6	M+I6	…	…
H	M+E+I7	M+E+I7	M+I7	M+I7	…	…

Figure 2.24. General microplate layout in the ligand displacement (IC_{50}) experiments (96 well microplate). Enzyme (E), binder (I) and mixture containing fluorophore and TX-100 (M). All dilutions in 13 mM HEPES buffer (pH 7.0)

All experiments included controls. Controls in the binding affinity experiments included only fluorophore, no enzyme and thus fluorescence intensity remained high. Negative controls in the ligand displacement experiments included enzyme and fluorophore (low fluorescence intensity) and positive controls contained a fluorophore in each measured nucleotide sugar concentration in the absence of enzyme (high fluorescence intensity).

All raw data is normalised as a percentage of the control which possess the highest fluorescence intensity on the examined microplate. The results of the binding affinity assay are presented in excel by plotting the normalised fluorescence intensity over the range of various enzyme concentrations. The IC_{50} values were calculated by plotting the normalised fluorescence intensity over the range of sugar nucleotide concentrations to GraFit 7.0.3.

2.7.4 Assay parameters; reaction volume, concentration of fluorophore and DMSO, buffer and incubation time

Linearity of the fluorophore was investigated in the range of 39 nM to 2500 nM at the total reaction volume of 100 µL and 200 µL. All dilutions were prepared in 13 mM HEPES pH 7.0 buffer.

Comparison of Tris/HCl and HEPES buffers were examined at pH 7.4 to 8.3 and 6.8 to 7.8 respectively, the impact of the DMSO concentrations was examined at 0, 2, 5 and 10 % (v/v) and final incubation times of 10, 20, 30 and 60 minutes were examined. Three different pre-incubations were performed:

- Incubation 1: UDP and fluorophore incubated for 10 minutes then LgtC added
- Incubation 2: UDP and LgtC incubated for 10 min minutes then fluorophore added
- Incubation 3: LgtC and fluorophore incubated for 10 minutes then UDP added

The final reading was taken 20 minutes after the last component was added. All experiments were performed with the LDA with UDP, UDP-Gal, or both, in various concentrations (0.01 to 10000 μM).

2.7.5 Activation of LgtC

Activation of LgtC was investigated by binding affinity assay. LgtC was activated in 5 mM of DTT in 13 mM HEPES buffer (pH 7.0) for 30 minutes at 37 °C, prior to use. The highest final LgtC concentration in the well was 32 μM and diluted by half, down to the lowest concentration of 0.004 μM.

2.7.6 Influence of metal

The experiment was performed in the absence and in the presence of 10 mM $MnCl_2$. A $MnCl_2$ solution in 13 mM HEPES buffer (pH 7.0) was prepared daily, prior to experiments. Binding affinity experiments included the highest final LgtC concentration in the well of 27 μM, and diluted by half down to the lowest concentration of 0.42 μM. The influence of the metal was also investigated in two separate ligand displacement assays with UDP and UDP-Gal at concentrations from 0.01 to 10000 μM.

2.7.7 Non-specific binding and influence of surfactant

Non-specific binding was investigated by the binding affinity assay in the absence and in the presence of BSA (Mw 66 kDa), and in 0 %, 0.001 % and 0.01 % of TX-100 in the absence of BSA. The highest final BSA/LgtC concentration in the well was 8.2 μM and diluted by half, down to the lowest concentration of 0.002 μM. The experiment in the presence of various

concentrations of TX-100 was titrated with LgtC concentrations from 0.008 to 64 µM. UDP concentrations in the LDA were from 0.001 to and 10000 µM.

2.7.8 Influence of LgtC concentration

Various concentrations of LgtC were investigated by the LDA at concentrations of 1.0, 5.0 or 15 µM in the presence of TX-100 (0.01 %) and DMSO (10 %) (v/v). UDP concentrations varied from 0.005 to 5000 µM.

2.7.9 Reproducibility experiment

Reproducibility was determined on the 96 well microplates and the experiment was performed on two different days. The negative control (low fluorescence intensity) included 5 µM LgtC and 200 nM fluorophore, the positive control (high fluorescence intensity) included 5 µM LgtC, 200 nM fluorophore and 5.0 mM UDP. The control (high fluorescence intensity) included 200 nM fluorophore. The experiment was performed in the absence and in the presence of TX-100 at 0.01 %. Each Z'-factor plate contained 32 replicate negative controls, 32 replicate positive controls and 32 control wells (Figure 2.25). On the second day column one was positive control (P), column 2 was control (C) and column 3 was negative control (N) etc.

	1	2	3	4	5	6	7-12
A	N	P	C	N	P	C	Etc.
B	N	P	C	N	P	C	Etc.
C	N	P	C	N	P	C	Etc.
D	N	P	C	N	P	C	Etc.
E	N	P	C	N	P	C	Etc.
F	N	P	C	N	P	C	Etc.
G	N	P	C	N	P	C	Etc.
H	N	P	C	N	P	C	Etc.

Figure 2.25. Microplate layout in the reproducibility experiment. Negative control (N), positive control (P) and control (C)

2.7.10 Stability of LgtC

The same LgtC stock (321 µM) was used in the stability experiments. The experiments were performed in identical conditions two and five months after expression. The specific activity was determined directly after and five months after the expression.

The specific activity was determined by NADH coupled assay (conditions described in Chapter 2.7.2). The turnover of the substrates to the products were monitored at 0.54, 5.4 and at 54 µM of LgtC.

The binding affinity was performed on the 96 well microplates, at fixed fluorophore concentration (200nM) and variable LgtC concentrations from 0.013 µM to 13 µM.

The LDA was performed on the 96 well microplates, at fixed fluorophore and LgtC concentrations, 200nM and 4 µM respectively. UDP-Gal concentrations vary from 0.01 to 10000 µM.

The K_m experiments were performed with phosphatase coupled assay (described in Chapter 3.4 and the reaction conditions in chapter 3.6.3.1). The experiments were carried out in mixture of $MnCl_2$ (5.0 mM), CEL (1.0 mg/ mL), CIP (10 U/mL) and lactose (2.0 mM) in 13 mM HEPES buffer (pH 7.0). LgtC was activated and the experiment were carried out at 20 nM and UDP-Gal (donor substrate) at 3.1, 6.3, 12.5, 25, 50, 100, 200 µM. The K_m values were calculated by the initial velocity over the range of UDP-Gal concentrations to GraFit 7.0.3.

2.7.11 Assay adaptation to 384 well microplates

Nunc and Greiner non-coated 384 well microplates were compared by performing fluorophore calibration curve from 12.5 to 1000 nM (ten points) in various reaction volumes (10 µL for Nunc, 15 µL for Greiner, 20, 30 40 and 50 µL). Both, the binding affinity and the ligand displacement assays were performed at 20 µL in triplicate in the presence of TX-100 (0.01 % (v/v)). The ligand displacement assay was performed with UDP-Gal, UDP, UMP and uridine at concentrations from 0.005 to 5000 µM.

2.7.12 Investigation of other GTs

NGT and TcdB were received from the collaborators. Dr Jon Cuccui and Professor Brendan Wren provided NGT, and Professor Klaus Aktories and Dr Thomas Jank provided TcdB. LgtC and β1,4-GalT were expressed and purified in house. The experiments were carried out in triplicate on 384 well microplates with increasing protein concentration, at 200 nM of fluorophore and at 0.01 % of TX-100 (v/v). Reaction conditions are listed in Table 2.13.

Table 2.13. Conditions used in binding affinity assay for studied GTs

Enzyme	Buffer	[Metal]	pH	Fluorophore
LgtC	13 mM HEPES	-	7.0	5FTUDP-Gal
β1.4-GalT	13 mM HEPES + 50 mM KCl	5 mM $MnCl_2$	7.0	5FTUDP-Gal
TcdB	25 mM Tris, 150 mM NaCl, 150 mM K_2SO_4	5 mM $MgCl_2$ and 5 mM $MnCl_2$	7.5	5FTUDP-Glc
NGT	300 mM NaCl + 250 mM imidazole	-	8.0	5FTUDP-Gal and 5FTUDP-Glc

NGT has a low binding affinity to UDP-Gal,[141] therefore, the experiment was carried out also with 5FTUDP-Gal. However, the fluorescence quenching was not observed. The binding affinity experiment with NGT was also attempted in HEPES buffer. The fluorescence was quenched linearly with increasing concentration of the pure imidazole buffer (NGT storing buffer) in HEPES buffer, therefore, the assay was performed in the imidazole buffer to prevent non-specific quenching.

3 Biochemical assays

In the previous chapter (Chapter 2), a ligand displacement assay (LDA) was developed for the identification of glycosyltransferase (GT) inhibitors. The LDA monitor only the displacement of the fluorescence donor analogue, whereas biochemical assays are designed for monitoring the enzymatic turnover. As the LDA is unsuitable for the determination of enzymatic activity the biochemical assay was required. The aim of this chapter was to develop a biochemical assay in order to determinate the kinetic parameters of LgtC and to evaluate inhibitors identified from compound screening (Chapter 4).

3.1 Introduction

A variety of biochemical assays have been developed for the characterisation of GTs.[63, 66] Ideally the enzymatic reaction is monitored in real time and monitoring the glycosylated product formation is favourable because small concentration changes in the product formation are easier to detect than the concentration changes in donor substrate concentrations. As the natural acceptors are poor light emitters, the development of biochemical assays based on monitoring the glycosylated product can be problematic. Radiochemical and mass spectrometry assays[66, 80, 158] have been reported for monitoring the formation of glycosylated products however, these approaches have disadvantages (i.e. radiochemical waste and expensive equipment). Other techniques such as detecting the post-derivated product[72, 74] and labelled acceptors[73, 159] are reported. Natural GT reactions can be monitored by detecting the UV-active nucleobase of the donor substrate or the secondary product by UV-detection.[63] The assays based on the chromatographic separation by HPLC (high-performance liquid chromatography) coupled with UV-detection[70] are suitable for detection both the depletion of NDP-donor or formation of NDP-product simultaneously. In addition, the essential determination of the amount of enzymatic hydrolysis of the donor is possible under the assay conditions. Many other GT assays have been developed based on monitoring the secondary product with various detection methods such as fluorescent chemosensors[92, 160, 161] and coupled enzymatic assays.[84, 85]

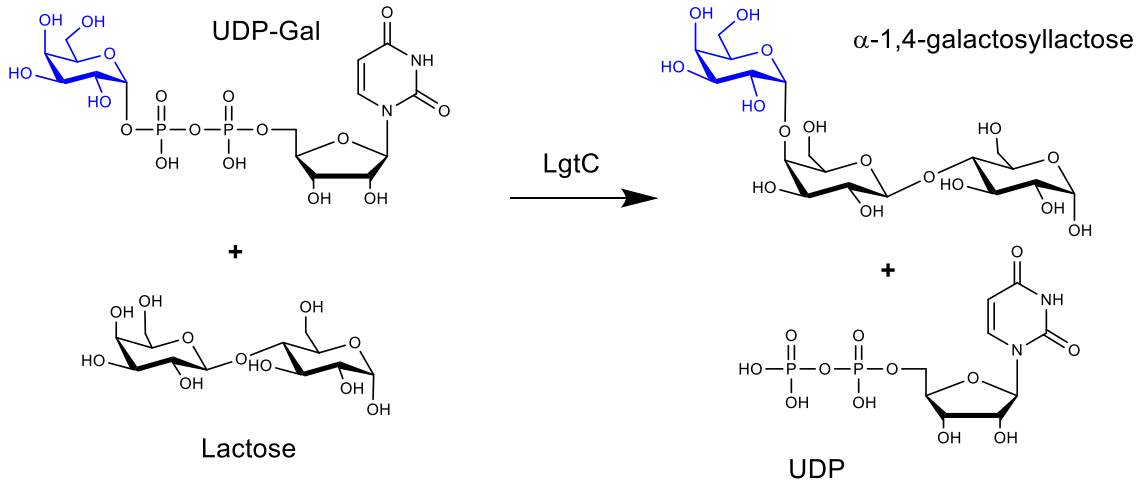
The aim of the chapter was to develop a biochemical assay in order determine enzyme kinetic parameters catalysed by LgtC and to evaluate inhibitors from the compound library screening (Chapter 4). Several biochemical assays were explored:

- HPLC method was optimised for the separation of NDP donor and product of LgtC catalysed reaction, and the HPLC method was adapted for determination of the kinetic parameters of the LgtC catalysed reaction
- A novel GT assay was introduced. The development progress was monitored alongside the HPLC-based method. During the assay development, unexpected features of LgtC were observed. LC-MS/MS method was then developed and utilised for further investigations
- A phosphatase coupled assay was successfully optimised for the monitoring the LgtC catalysed reaction

3.2 HPLC method

HPLC is a generic analytical method for separation a mixture of compounds, compound identification and quantitative analysis. Several chromatographic techniques are reported for the analysis of nucleotides and sugar nucleotide simultaneously from complex biological matrices.[162] Ion-exchange chromatography (IEC) and ion-pair chromatography (IPC) are the most popular techniques to separate polar and ionic molecules such as nucleotides and nucleotide sugars.[163-167] IPC is a general technique to improve the retention of polar compounds in reversed phase chromatography (RP-LC). Ion-pair reagents that have a charge opposite to the molecule of interest are added to the mobile phase. The improvement of the retention can occur either by dynamic ion-exchange on the stationary phase or formation of neutral ion-pairs which are retained in non-polar stationary phase.[168]

HPLC method can be easily adapted for the determination of kinetic parameters of enzymatic reactions. Saccharides are non- UV-active, therefore, monitoring the primary product (galactosyllactose) of LgtC catalysed reaction is impossible without derivatisation (Scheme 3.1). The donor sugar (UDP-Gal) and the secondary reaction product (UDP) reaction are UV-active and they are possible to quantify by HPLC coupled with UV-detector. In order to obtain kinetic properties of LgtC, HPLC-based biochemical assay was developed.



Scheme 3.1. LgtC catalysed reaction can be monitored by HPLC/UV due to the UV-active catalytic base of UDP

3.2.1 Method optimisation

The intention of the HPLC method was to separate and quantify the depletion of UDP-Gal and formation of UDP in LgtC catalysed reaction. Formation of the side product (UMP) was expected by enzymatic or chemical hydrolysis of UDP-Gal, therefore, UMP was also included to the method development. Before the HPLC-based method was adapted for monitoring the LgtC catalysed reaction, the retention of the compounds and HPLC parameters were optimised using standard (UMP, UDP and UDP-Gal) solutions.

Ion-pair chromatography has been proved to be an effective separation and quantification method for polar compounds such as nucleotides and nucleotide sugars, including UDP-Gal and UDP.[70, 162, 169, 170] Kochanowski and co-workers developed a sensitive and reproducible IPC method for the identification and quantification of intracellular nucleotide and nucleotide sugar levels from Chinese hamster ovary cells.[169] Because UDP-Gal, UDP and UMP were separated with this chromatographic method, the method was selected to optimise the analysis of the components of the LgtC catalysed reaction.

Kochanowski separated 8 nucleotides and 5 sugar nucleotide by using a standard C18 column (3 µm particle size, 15 cm x 4.6 mm).[169] The eluent system consisted of two buffers. Buffer A was 100 mM potassium phosphate buffer including 8 mM tetrabutylammonium hydrogen sulphate (TBAHS) as an ion-pair reagent and buffer B was 70% buffer A with 30% methanol (MeOH). The retention time of UDP-gal and UDP was less than 15 minutes and the total run time was 47 minutes, therefore, the gradient of the run was modified to obtain shorter analysis time: 0 to 77 % buffer B in 27 minutes was changed to 0 to 50 % buffer B in 15 minutes (Table 3.1).

Table 3.1. Gradient steps of HPLC method for analysis of UMP, UDP-Gal and UDP

Time (min)	Buffer A (%)	Buffer B (%)
5	100	0
15	100-50	0-50
2	50	50
1	50-100	50-0
5	100	0

A mixture of UMP, UDP and UDP-Gal (25 µM each) was analysed by the optimised HPLC gradient at flow rate 1 mL/min and detected at 265 nm. The column temperature was set to 30 °C and the injection volume was set to 50 µL.

All compounds were retained on the column and the peaks were clearly separated with the increasing gradient of MeOH concentration (Figure 3.1). The peak shape and response of all components were satisfactory at the injection volume 50 µL. The peaks were identified by comparing the retention times of individual standards. UMP, UDP-Gal and UDP peaks retained at 5.8, 9.4 and 12.8 minutes respectively.

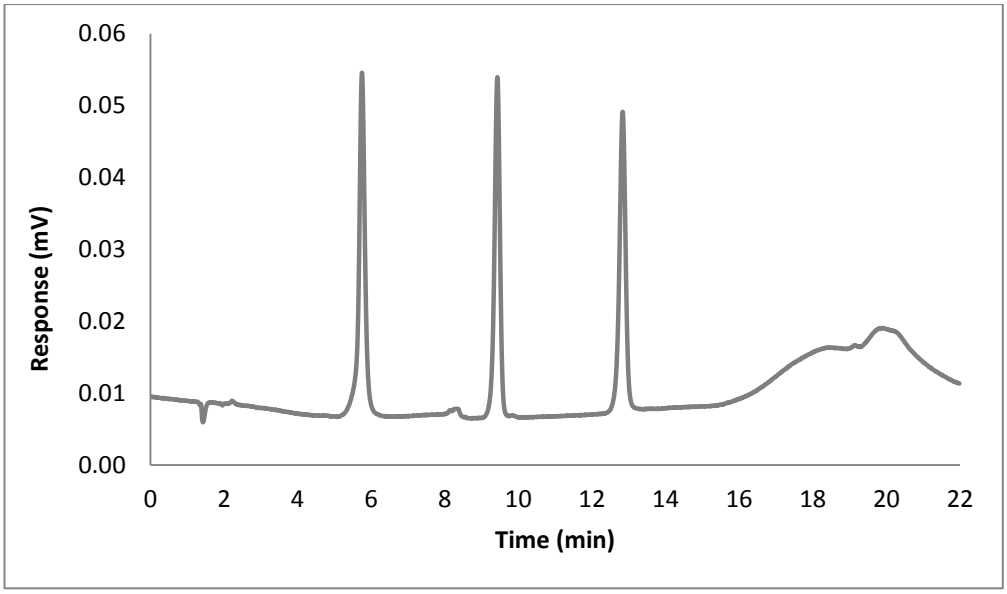


Figure 3.1. UMP, UDP-Gal and UDP eluted at 5.8, 9.4 and 12.8 min, respectively. Separations were performed with the optimised gradient (Table 3.1), an RP column and the wavelength of UV detection was 265 nm

3.2.2 System suitability

The system suitability test was performed to ensure that the complete testing system including the instrument, the used reagents and the column were suitable for the intended application.[171] The consistency of the system performance (precision of replicate injections) and the chromatographic suitability (e.g. tailing factor, column efficiency and resolution) are the main components of the system suitability. The intention of the HPLC method was to quantify the concentration of UDP-Gal and UDP during the enzymatic reaction. The following system suitability calculations were performed to ensure that the optimised HPLC method was adequate for particular analysis.

Precision measurements define how reproducible the results are and assure that the autosampler is delivering the same volume each time. The precision is determined with measurements of multiple sampling of the same homogenous sample. The precision is expressed as the standard relative deviation (%RSD)[171] calculated as in Equation 2, where St is standard deviation and M is mean.

Equation 2

$$\%RSD = \frac{St \times 100\%}{M}$$

The column efficiency (N) is a measure of the dispersion of a peak. The column efficiency depends on the retention time (RT) and peak width ($W_{1/2}$) at half height of the peak. The higher the retention time and the narrower peaks, the greater the column efficiency and the more effective separation.[171] The column efficiency was calculated as in Equation 3.

Equation 3

$$N = 5.45\left(\frac{RT}{W_{1/2}}\right)^2$$

The resolution (R) illustrates how well two peaks are separated. The resolution of the two peaks can be calculated using the retention times ($(RT)_A$ and $(RT)_B$) and the peak width at the base (W_A and W_B). When the resolution is > 1.5 the valley between two peaks returns to the baseline.[171] Resolution was calculated as in Equation 4.

Equation 4

$$R = \frac{2(RT)_B - (RT)_A}{W_A + W_B}$$

The tailing factor (T) is a measure of the symmetry of a peak that can be calculated from the peak width at 5% height ($W_{0.05}$) and the distance from peak front to apex point at 5% height (*f*). Tailing factor of a perfectly symmetrical peak corresponds to value 1.[171] Tailing factor was calculated as in Equation 5.

Equation 5

$$T = \frac{W_{0.05}}{2f}$$

The precision of the retention time, peak area and height were determined for UMP, UDP-Gal and UDP at 25 µM by performing six replicate injections, and calculating the %RSD. UDP-Gal may be hydrolysed to UMP and inorganic phosphate via UDP, thus UMP was added to the precision measurements. Column efficiency and tailing factor of the peak were determined for UDP-Gal and UDP at 25 µM. In addition, the linearity of UDP-Gal and UDP was determined as it was important to obtain an accurate concentration of donor sugar and product of enzymatic reaction. The linearity of a method is defined as its ability to obtain test results that are directly proportional to the sample concentration within a given range.[171] The linearity was studied in a concentration range from 1.0 µM to 200 µM of UDP-Gal and UDP,

and the linear regression (R^2) was determined by plotting the peak area versus the concentration of measured sample. All the system suitability results are presented in Table 3.2.

Table 3.2. System suitability tests determined at 25 µM of UMP, UDP-Gal and UDP. %RSD is in brackets

Component	Retention time (min)	Peak area (uV/sec)	Peak height (uV)	Column efficiency (N)	Tailing factor (T)	Linearity[a] (R^2)
UMP	5.76 (0.1)	466542 (0.2)	50522 (1.2)	-	-	-
UDP-Gal	9.24 (0.1)	297493 (0.4)	33350 (0.5)	22667	1.1	0.997
UDP	12.83 (0.04)	497848 (0.6)	41708 (0.5)	25667	0.9	0.996

[a]Determined at range of 1.0 µM to 200 µM

UMP retained first then UDP-Gal and finally UDP, by the optimised HPLC method. All nucleotides and sugar nucleotides were visibly separated except a small apparent shoulder of UDP-Gal. Further investigations showed that the impurity was derived from UDP-Gal (97 % pure by the manufacturer). The resolution between UDP-Gal and the impurity peak was 1.6. Resolution higher than 1.5 is satisfactory for an accurate integration because the valley between two peaks touches to the baseline.[171] Based on visual evaluation, the limit of quantification (LOQ) of UDP-Gal and UDP was approximately 1.0 µM. Below the concentration of 1.0 µM the peaks were not clearly visible and the peak integration was imprecise.

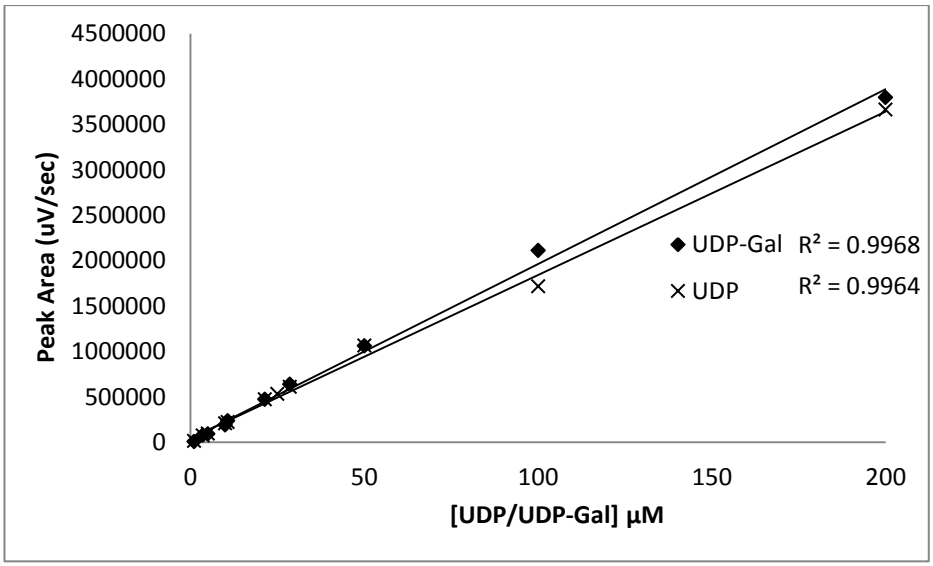


Figure 3.2. Peak area increased linearly with increasing concentration of UDP-Gal and UDP in range of 1.0 µM to 200 µM with optimised HPLC method

The recommended %RSD for precision is > 1 % by U.S. Food and Drug Administration (FDA).[172] Based on the recommendations of FDA the reproducibility of retention time, height and peak area were excellent for all components except for UMP. %RSD of the peak height of UMP was slightly out of range (1.2 %). This can cause variable results if peak height of UMP is used for the determining the concentration of UMP. However, this result was not significant because the intention was to record peak area of UDP-Gal or UDP accurately. UDP and UDP-Gal peaks were symmetrical and the peaks are clearly separated from the baseline. The tailing factor and the column efficiency were in the limits of FDA recommendations; ≤ 2.0 and > 2000 respectively. In addition, the linear regression (>0.996) indicated an excellent relationship between peak area and concentration of UDP-Gal and UDP in range of 1.0 µM to 200 µM (Figure 3.2).

3.2.3 Adaptation of the HPLC method for monitoring LgtC catalysed reaction

The HPLC method was developed to quantify UDP-Gal and UDP, therefore, the HPLC method was adapted to follow the LgtC catalysed reaction by monitoring the depletion of UDP-Gal or formation of UDP. The suitability of the optimised HPLC-based method was evaluated by determining repeatability (%RSD) and K_m of UDP-Gal of the LgtC catalysed reaction in order to use the assay for compound characterisation.

The reaction conditions used by previous laboratory members were adapted to monitor the LgtC catalysed reaction.[52] The reactions were carried out in 50 mM MOPS buffer, including 10 mg/mL of BSA and 20 mM of $MnCl_2$ at pH 7.0. The lactose concentration was at 2.0 mM, UDP-Gal was fixed or variable depending on the experiment, and reactions were initiated by adding LgtC. After 10 minutes incubation at 37 °C the reactions were stopped by placing the samples in dry ice. The samples were then thawed prior to the injection and the optimised HPLC method was used to quantify formation of UDP or depletion of UDP-Gal.

Previous results (chapter 2.2.2) suggested that LgtC activation is not necessary when binding affinity is monitored. However, enzyme kinetic studies have suggested that activation with DTT increases catalytic activity of LgtC.[147] Therefore, the effect of the activation on LgtC activity was studied by HPLC-based assay. UDP formation was monitored in various activated and non-activated LgtC concentrations (Figure 3.3). The results showed a 1.4-1.9 fold

increase in LgtC activity, supporting the hypothesis that the activation only effects the acceptor binding. The results suggested that activation creates a favourable conformational change for the acceptor binding and product release. Therefore, LgtC was activated prior to use for all future experiments.

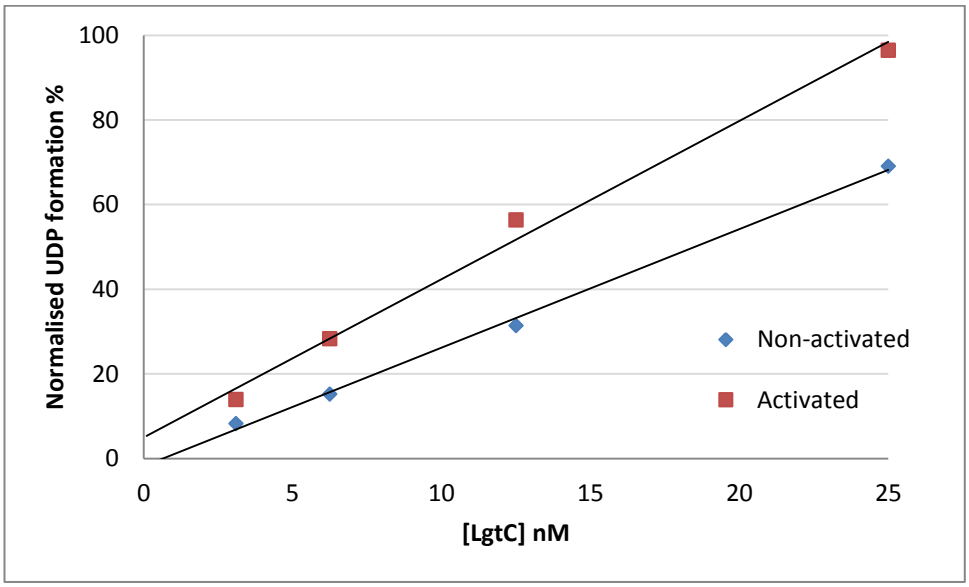


Figure 3.3. UDP formation of non-activated versus activated LgtC catalysed reaction followed by HPLC-based method with detection wavelength 265 nm. Lactose and UDP-Gal concentrations 2.0 mM and 200 µM, respectively

Reproducibility of HPLC-based method was studied by performing LgtC catalysed reaction eight times in identical conditions over one day. The UDP-Gal concentration was fixed to 60 µM and reactions were stopped after 10 minutes incubation by placing the samples in dry ice. Formation of UDP peak was monitored and the variability was 4.1 %RSD (Table 3.3).

Table 3.3. Reproducibility of LgtC catalysed reaction by monitoring UDP peak area (*n*= 8)

Peak area of UDP (uV/sec)	539570
STDEV	21889
%RSD	4.1

[a]Experiment performed at 2.0 mM of lactose, 0.63 nM of activated LgtC and at 60 µM UDP-Gal

The suitability of the optimised HPLC method for enzymatic kinetic studies was evaluated by determining the K_m and K_{cat} of LgtC catalysed reaction (Table 3.4) by monitoring the formation of UDP. The initial reaction velocity was measured at various UDP-Gal concentrations ensuring the donor conversion to the product was 10 % or less. As demonstrated in previous chapter 2.2.1, the LDA was not affected by DMSO. Therefore, the

sensitivity of HPLC-based assay was examined in the presence of 10 % (v/v) DMSO ensuring capability to characterise possible hits from the compound library screening (Chapter 4).

Table 3.4. K_m and K_{cat} of UDP-Gal determined by the optimised HPLC-based method (*n*= 3)

Compound	K_m (µM)[a]	K_{cat} (s^{-1})
UDP-Gal	4.9 ± 1.6	6.0 ± 1.0
UDP-Gal (10 % DMSO)	2.2 ± 1.1	6.1 ± 3.0

[a]Experiment performed at 2.0 mM of lactose, 0.63 nM of activated LgtC and UDP-Gal concentrations of 1.6, 3.1, 6.3, 12.5, 25, 50, and 100 µM

The kinetic parameters of LgtC were determined by HPLC-based method. K_m was slightly lower than the literature values (13-29 µM[16, 108, 173]), however, the difference is probably from the conditions used in particular biochemical assay. The results indicated that DMSO did not have significant effect for the assay efficiency as K_m and K_{cat} remained on the same range. Low K_m value of UDP-Gal demonstrates a high binding affinity to LgtC which results that the used assay is required to detect accurately small substrate or product concentration changes. However, the accuracy of the obtained K_m values are questionable because the measurements required detection of low UDP concentrations near LOQ, particularly at 1.6 µM and 3.1 µM of UDP-Gal. Detection near the LOQ may cause inaccurate measurements which may significantly effect the initial velocity at the lowest UDP-Gal concentrations and can result in error on Michaelis-Menten curve. Thus, the slight difference between the K_m values in the presence and in the absence DMSO may be due to the poor detection sensitivity not the effect of DMSO.

Overall the HPLC-based method has disadvantages and advantages. A disadvantage is the sensitivity and high LOQ which complicate the accurate determination of small concentration changes of UDP and UDP-Gal in the enzymatic kinetic experiments. Therefore, the HPLC-based method may produce inaccurate results if it is used for the characterisation of potential LgtC inhibitors at low UDP-Gal concentrations. However, the HPLC- based method can be used for less efficient enzymes (higher K_m value) than LgtC. Additionally, a disadvantage of the HPLC-based method is the time consuming assay protocol which was seen during K_m determination. The time consuming steps are the equipment set up, the samples must be thawed prior to each injection and only one sample can be analysed at once. An advantage of the HPLC-based method is that it can be used for the direct quantification of the depletion of the donor substrate or formation of the product and the results can be used for producing Michaelis-Menten curves without the need for labels or

additional derivations. In addition, the hydrolysis product, UMP, can be monitored simultaneously with UDP-Gal and UDP. The HPLC-based method is not practical to perform or sensitive enough for the determination of kinetic parameters of LgtC. However, the method may generate useful information if unforeseen results are discovered with spectrophotometric assays where usually only one substance can be monitored.

3.3 Development of glycosidase coupled assay

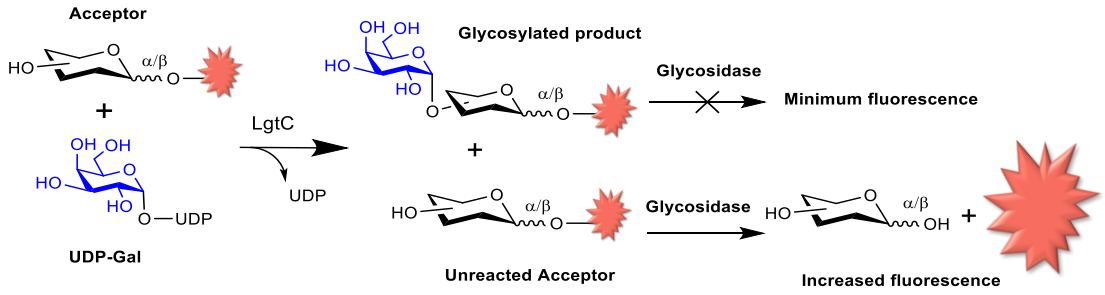
As described in the previous section, the HPLC-based assay was a time consuming assay protocol and the sensitivity was low (LOQ = 1.0 µM). Therefore, an alternative biochemical assay was required because the HPLC-based assay was impractical for the determination of enzyme kinetic values and for the evaluation of inhibitors identified from compound library screening.

Several GTs have been shown to use unnatural substrates.[174-177] For example, α1,3-GalT which naturally catalyses the transfer of galactose from UDP-Gal to disaccharides such as terminal lactose and *N*-acetyllactosamine containing glycoconjugates can use *p*-nitrophenyl-β-galactoside (*p*NPβGal) as an acceptor.[13] When *p*NPβGal was used as an acceptor, only α(1,3)-linked disaccharide derivative was formed, whereas in the reaction with free galactose as an acceptor, a mixture of regioisomers were formed.[13] In addition, Jamaluddin and co-workers supported the theory of acceptor substrate promiscuity of α1,3-GalT by performing a crystal structure of *p*NPβGal complex with α1,3-GalT and demonstrating that *p*NPβGal binds similarly to the natural acceptor substrate, *N*-acetyllactosamine.[178] Recent substrate engineering studies have shown that LgtC is able to use both, unnatural donor and acceptor substrates. LgtC has been shown to use α-galactosyl fluoride[174] and 2,4-dinitrophenyl-β-D-galactoside[13] as donor substrates in the presence of natural nucleotide, UDP. Systematic substrate engineering studies demonstrated an acceptor diversity of LgtC. LgtC was able to process variable unnatural substrates and depending on the alkyl or aryl substituent of the acceptor sugar, galactose was transferred via α(1-2), α(1-3) or α(1-4) linkages at synthetically useful rates.[12]

The substrate flexibility can be beneficial for the designing of biochemical assays of GTs because the GT substrates which are generally poor light emitters could be replaced with fluorescent, unnatural substrates. The sugar moiety of the acceptor molecule is important for biological activity and recent studies suggested that LgtC could utilise variable acceptor molecules with the requirement for the presence of sugar. Therefore, a novel fluorescence based biochemical assay was designed with LgtC. An advantage of this novel fluorescence assay is that it is sensitive and it can be performed on 96 well microplates.

3.3.1 Principle of the glycosidase coupled assay

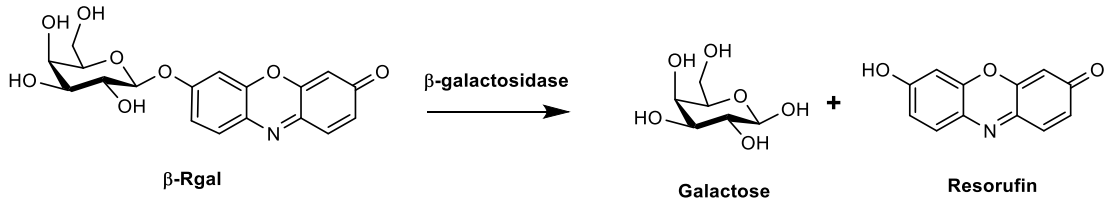
The aim was to develop a novel fluorescence-based biochemical assay by monitoring the LgtC catalysed reaction by using an unnatural fluorogenic acceptor coupled with a glycosidase that cleaves the sugar from the aglycone causing a change in fluorescence. Hypothetically, glycosylation might protect adjacent cleavage site from hydrolysis. The new assay was then designed by coupling the glycosidase and GT reaction. In this approach, a fluorogenic acceptor is subjected to glycosylation and then treated with a glycosidase that discriminates between the glycosylated and non-glycosylated fluorogenic acceptor molecules (Scheme 3.2). The resulting fluorescence signal is measured to determinate the amount of hydrolysis which is related to the degree of glycosylation. For example, in the presence of a high affinity inhibitor, the non-glycosylated fluorogenic substrate is hydrolysed by the glycosidase, leading in the increase of the fluorescence signal.



Scheme 3.2. General principle of the glycosidase coupled assay. LgtC catalyses the transfer of galactose from UDP-Gal to the fluorogenic acceptor producing glycosylated product which cannot be hydrolysed by specific glycosidase and fluorescence remains low. The unreacted acceptor is then hydrolysed by glycosidase resulting in high fluorescence

The requirements of this glycosidase coupled assay were (1) identify a fluorogenic/glycosidase pair, (2) fluorogenic substrate must be a good acceptor for LgtC and (3) there must be a large difference in fluorescence intensity signal between the glycosylated and non-glycosylated products. Exo-acting glycosidases[179] cleave a sugar at the non-reducing end of the chain and β-mode exo-acting glycosidases were the option with most potential for the coupling with LgtC catalysed reaction because they are unable to recognise α-linked sugars. Thus, the α1,4-linked galactose of the glycosylated product by LgtC is protected.

Resorufin modified sugars were of interest as fluorogenic acceptors because they have a long fluorescence emission wavelength that is less prone to interfere with fluorescent compounds in the characterisation of potential inhibitors, and additionally, the hydrolysed product (resorufin) is strongly fluorescent compared to the resorufin-sugars.[180] Modified resorufin molecules have been used for studying glycosidases such as cellulose,[181] mannosidase,[182] glucosidase[180, 183] and galactosidase[99] activities in various assay modes, including HTS.[184] A substrate of β-galactosidase, resorufin-β-galactopyranoside (Rgal)[185, 186] was a potential fluorogenic acceptor substrate for LgtC as it could mimic the natural acceptor (lactose).



Scheme 3.3. Resorufin-β-galactopyranoside (Rgal) hydrolysed by β-galactosidase results strongly in fluorescent resorufin (λ_{em} =590 nm, λ_{ex} =570 nm)

Rgal and β-galactoside "pair" was selected for the development of a coupled assay together with LgtC. The aim was to establish reaction conditions to detect resorufin which was comparable to the amount of unreacted Rgal. The designed assay consists of two separate reactions: β-galactosidase catalysed reaction and LgtC catalysed reaction. This coupled assay development contained the following studies:

- Linearity and stability of resorufin
- Optimisation of concentration of β-galactosidase
- Functionality of Rgal as an acceptor for LgtC by HPLC
- Optimisation of concentrations of LgtC and Rgal
- Positive and negative control investigation
- Optimisation of incubation time

3.3.2 Optimising β-galactosidase catalysed reaction

The functionality of the β-galactosidase catalysed reaction was investigated. β-galactosidase has been reported to function in 50 mM HEPES pH 7.3 buffer including 1.0 mM $MgCl_2$.[187] This buffer with 1.0 mM $MnCl_2$ was chosen for the preliminary experiments because of its similarity with the LgtC assay buffer that was used in the HPLC-based method (chapter 3.2.3). The aim was to find an optimal β-galactosidase concentration that hydrolyses galactose from Rgal rapidly. The hydrolysis of Rgal was investigated in real time with various β-galactosidase concentrations (Figure 3.4, A). At the highest β-galactosidase concentration (8.2 U/mL) all Rgal was hydrolysed very rapidly and to reach the same level of hydrolysis at 0.13 U/mL of β-galactosidase, 20 minutes incubation was required. 0.5 U/mL of β-galactosidase was selected for future experiments because at this concentration 10 minutes incubation time was optimum to hydrolyse all Rgal and the consumption of large amounts of material was prevented.

The fluorescence linearity of resorufin, the hydrolysis product of Rgal, was investigated in various concentrations of Rgal in order to discover a suitable acceptor concentration. The linearity of fluorescence of resorufin decreased at concentrations over 20 µM. The linear range of detected resorufin was between 0.08-10 µM (Figure 3.4, B) and the acceptor, Rgal, had to be used in this concentration range.

In addition, the stability of resorufin was investigated by monitoring the fluorescence of standard solutions of resorufin in various concentrations. Photobleaching of resorufin was not observed during one hour incubation at 37 °C (Figure 3.4, C). Therefore, the high stability of resorufin would make a long incubation time possible.

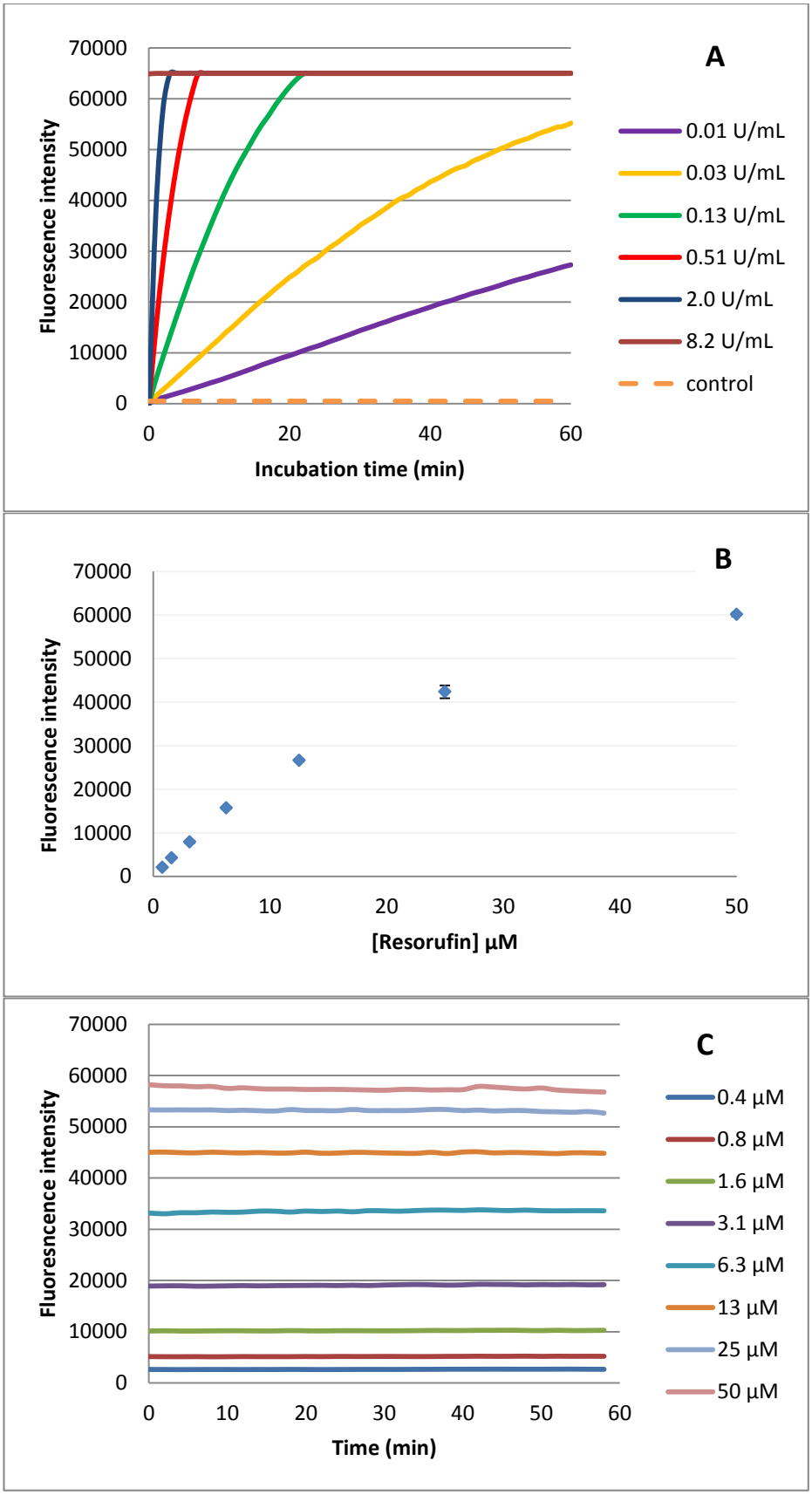


Figure 3.4. Hydrolysis of Rgal (20 µM) with various β-galactosidase concentrations (A). Fluorescence linearity of resorufin by hydrolysis of 0.8 to 50 µM Rgal by 0.5 U/mL of β-galactosidase (B). Stability of resorufin at concentrations of 0.4 to 50 µM (C). Fluorescence monitored at λ_{em} =590 nm and λ_{ex} =544 nm

3.3.3 Coupling LgtC reaction and β-galactosidase reactions

Preliminary tests were performed with the HPLC-based assay (Chapter 3.2.3) to evaluate Rgal functionality as an acceptor with LgtC. The LgtC catalysed reaction was performed by using Rgal as an acceptor instead of lactose. The results showed that the UDP-Gal peak decreased and UDP increased with increasing concentration of LgtC (Figure 3.5). This indicated that LgtC used Rgal as an acceptor.

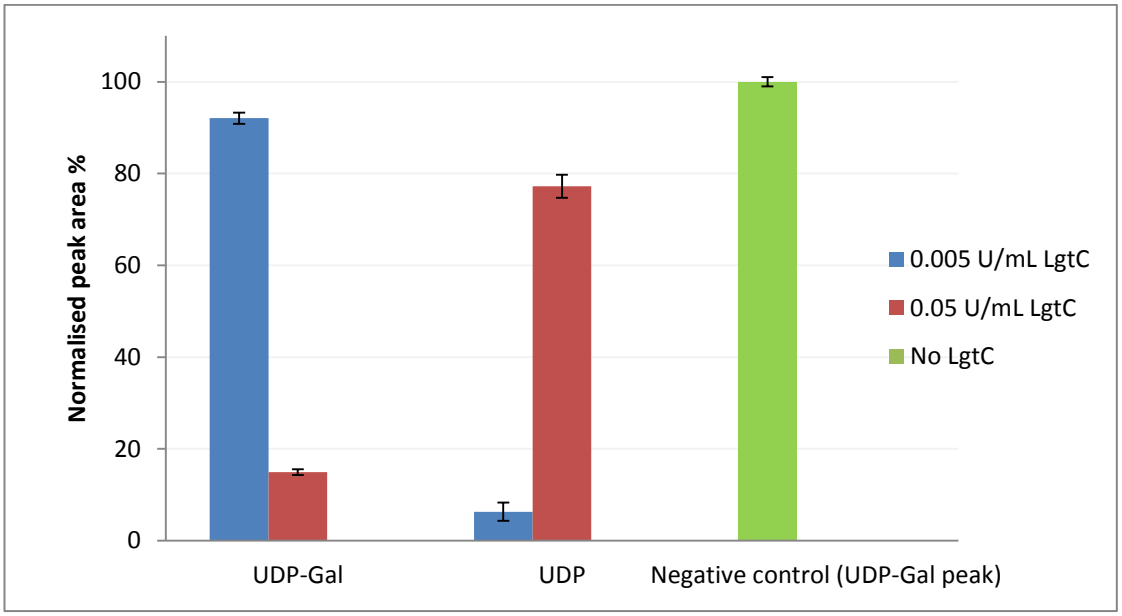


Figure 3.5. Reaction process of LgtC catalysed reaction monitored by optimised HPLC method (Chapter 3.2.1). The reaction monitored at 0.005 U/mL and at 0.05 U/mL of LgtC, and compared to the peak area of UDP-Gal in the negative control where LgtC was not added. Rgal at 100 µM was used as an acceptor and UDP-Gal at 60 µM as a donor.

The LgtC reaction and β-galactosidase reactions were then coupled in order to monitor the glycosylation by fluorescence detection. At the start, positive and negative controls were investigated. A positive control, when LgtC is active, results in a weak fluorescence signal as Rgal is galactosylated and protected from the β-galactosidase. The positive control was then performed without adding β-galactosidase to the reaction mixture. A negative control would result in a high fluorescence signal and would be observed in the absence of glycosylation (i.e. LgtC is non-active). The negative control was investigated in two ways: in the absence of LgtC and in the absence of UDP-Gal.

The fluorescence intensity was assumed to be the same for both negative controls. Initially the control experiments were performed with activated LgtC, resulting in stronger fluorescence intensity for the negative control in the absence of LgtC compared to the negative control in the absence of UDP-Gal (Figure 3.6). The experiment was repeated with non-activated LgtC and the fluorescence intensity of both negative controls reached the same level. These results indicated that possibly DTT quenches the fluorescence intensity of resorufin, therefore, subsequent experiments were performed without LgtC activation and the negative control was then performed in the absence of UDP-Gal. The control experiments showed a low fluorescence signal on the positive control, therefore, the positive control was performed in the absence of β-galactosidase.

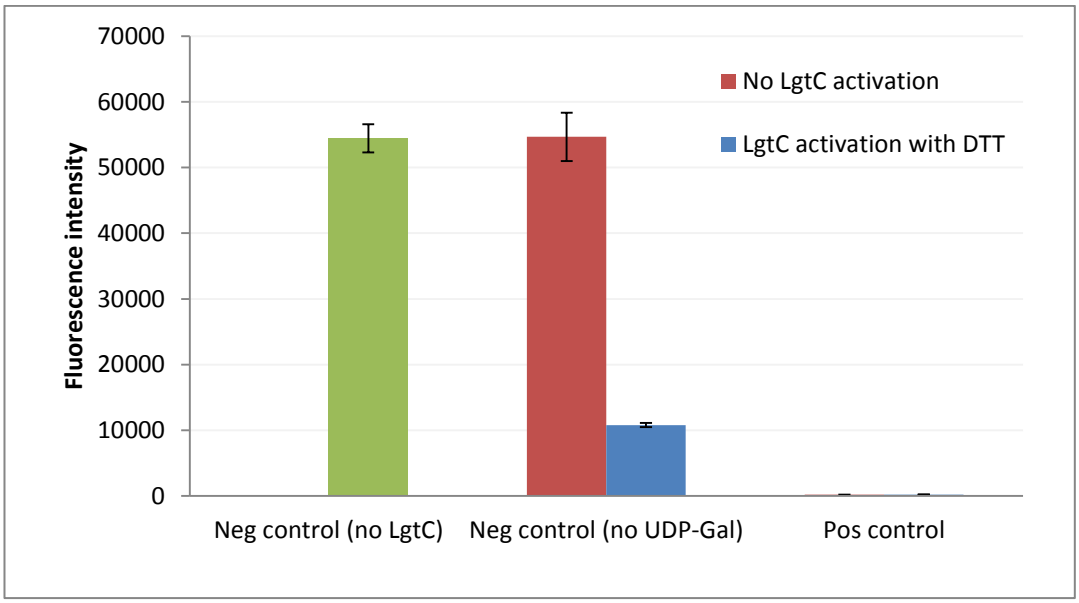


Figure 3.6. Negative and positive control investigations by coupling LgtC and β-galactosidase reactions. The experiment was performed at fixed concentrations of UDP-Gal (50 µM), Rgal (6.3 µM), LgtC (0.1 U/mL) and β-galactosidase (0.5 U/mL). Fluorescence monitored at λ_{em} =590 nm and λ_{ex} =544 nm

The reaction progress of the LgtC catalysed reaction was then investigated with various UDP-Gal concentrations and various incubation times, at fixed Rgal concentration. The assumption was that if the reaction is completed then almost all Rgal is glycosylated and a low fluorescence signal would be observed, similar to the fluorescence intensity level as the positive control. As illustrated in the Figure 3.7, the reaction progressed slowly with increasing incubation time at various UDP-Gal concentrations, more rapidly in higher concentration than lower concentrations. The fluorescence intensity remained high, indicating that only 40 % of the Rgal was glycosylated at 500 µM of UDP-Gal with the longest

incubation time. Therefore, these results indicated that the glycosylation was not completed in the reaction conditions. The utilised LgtC concentration was doubled in the experiment compared to the preliminary test with HPLC-based method (Figure 3.5). The HPLC experiment showed that nearly all UDP-Gal was consumed in the reaction, therefore, the utilised LgtC concentration was assumed to be sufficient to complete the glycosylation. However, these results suggest that at least 60 % of Rgal was free in the reaction mixture in the excess of UDP-Gal and after 60 minutes incubation.

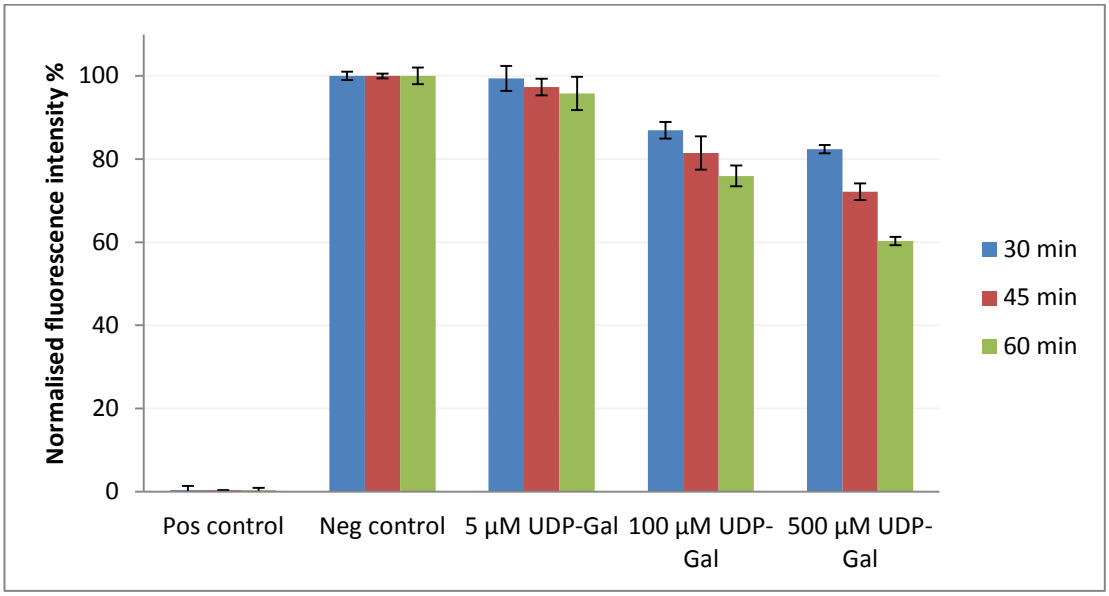


Figure 3.7. Glycosidase coupled reaction in various UDP-Gal concentrations and various incubation times. Reactions were carried out at 6.3 µM of Rgal, 0.1 U/mL LgtC and 0.5 U/mL β-galactosidase. Fluorescence monitored at λ_{em} =590 nm and λ_{ex} =544 nm

There are three possible explanations for the observed results: Rgal is a low affinity acceptor for LgtC, UDP-Gal is rapidly hydrolysed or LgtC transfers more than one galactose to the Rgal. The glycosidase coupled assay was developed to detect the unreacted Rgal without providing information about the reaction products or consumed UDP-Gal. To investigate further whether UDP-Gal was hydrolysed or Rgal was polygalactosylated, a method which could monitor both the consumption of the substrates and the formation of the products was required.

The HPLC based-method was appropriate for monitoring the reaction because all components were UV-active. The LgtC catalysed reaction was then adapted to a HPLC-based method. The concentration of UDP-Gal and Rgal were fixed and the reaction progress was

monitored at 30, 45 and 60 minutes incubation time. One new peak was observed on the HPLC trace when the product peaks of the LgtC catalysed reaction was analysed (Figure 3.8, A). The peak eluted at 15.2 minutes, after UDP and before Rgal, suggested the peak is glycosylated Rgal. A small UMP peak was observed but the peak size remained the same during the incubation time. As illustrated on the Figure 3.8 (B), after 60 minutes incubation, all UDP-Gal was consumed thus approximately 70 % of Rgal was unreacted. This indicated that LgtC possibly transferred more than one galactose to Rgal.

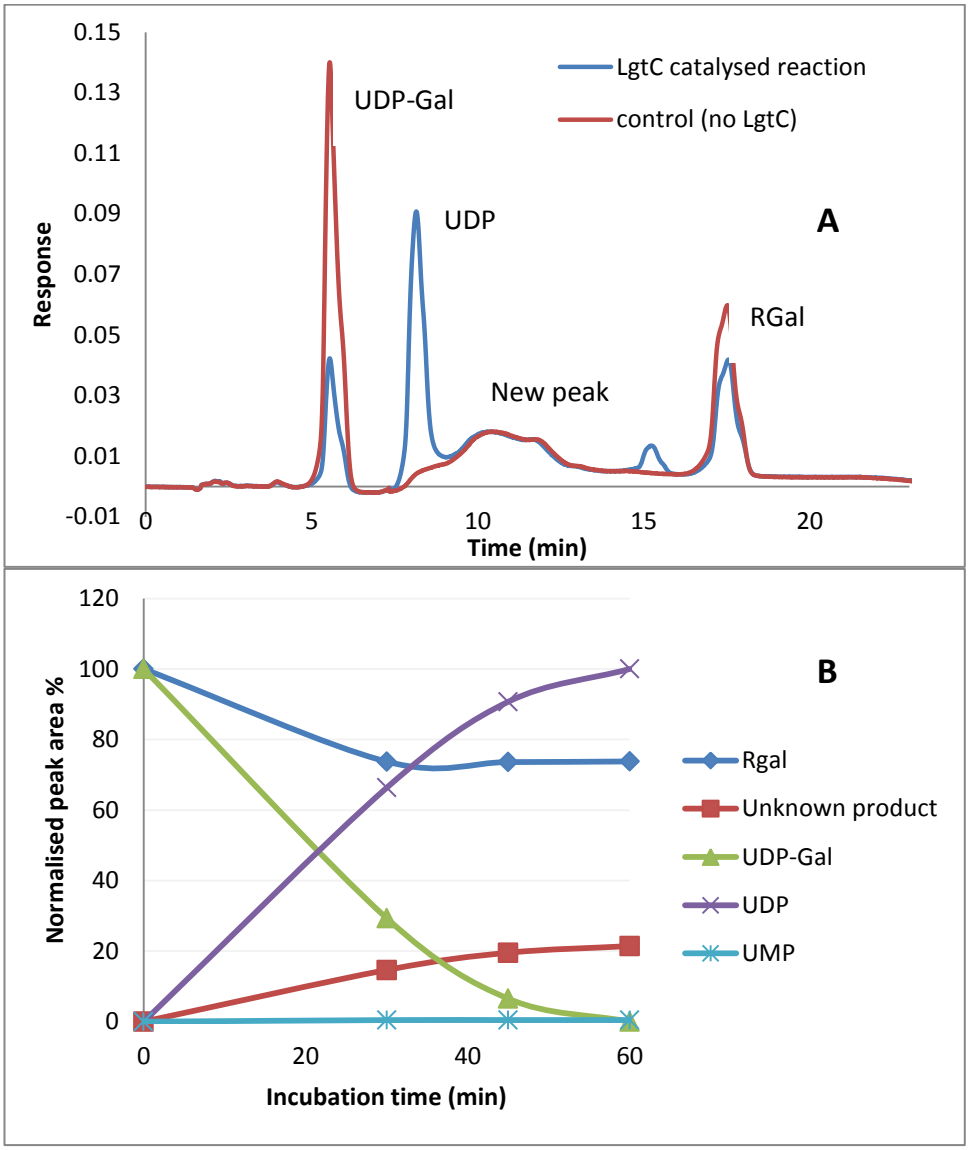


Figure 3.8. LgtC catalysed reaction stopped after 30 minutes incubation (blue trace) overlaid with the control, in the absence of LgtC (red trace) (A). LgtC catalysed reaction progress monitored by HPLC-based method (B). The reaction was carried out at 200 µM of UDP-Gal and Rgal, 0.02 U/mL of LgtC in 50 mM HEPES buffer including 1.0 mM $MnCl_2$ and $MgCl_2$ (pH 7.1). All components analysed at 265 nm

The Rgal was the limiting factor of the reaction rate because the concentration used was lower than the UDP-Gal concentration. However, the results obtained by the fluorescence based glycosidase coupled assay and the HPLC-based method, demonstrated that Rgal remained non-glycosylated at high UDP-Gal concentrations suggesting that UDP-Gal was not consumed in a one to one relationship to Rgal in the LgtC catalysed reaction. Therefore, additional investigations were required in order to identify the reaction product.

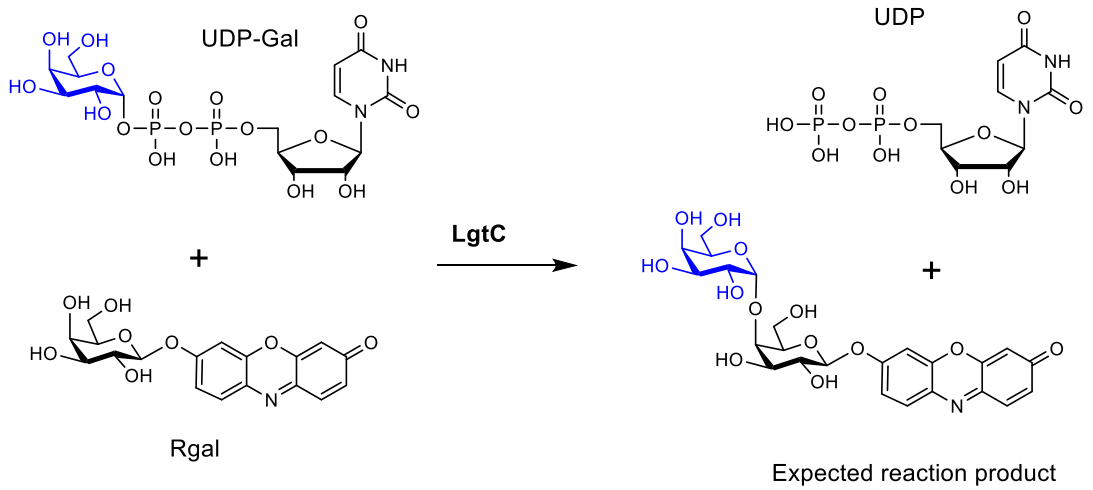
3.3.4 LCMS method development for the identification of the unknown product

Mass spectrometry (MS), which is a sensitive technique for the determination of molecular weight of molecules, was selected for the identification of the unknown product of the LgtC catalysed reaction when Rgal was used as an acceptor. MS is a widely used detection method and, combined with a separation method such as liquid chromatography (LC) allows an analysis of complex samples.[188] Connection of the two is not trivial, since the mass spectrometry requires that the analyte must be ionised in the gas phase. Ionisation methods such as electrospray ionisation (ESI),[189] allow for the transfer of large molecules from the liquid phase to the gas phase without breaking and if needed, also ionise them. Previously, the HPLC-based method was successfully adapted to monitor the donor and product of the LgtC catalysed reaction (Chapter 3.2.1). However, this HPLC method was unsuitable with the ESI/MS detection because a non-volatile ion-pair reagent was used. Therefore, an alternative chromatographic method was required for the separation of all reaction components.

Hydrophilic interaction liquid chromatography (HILIC) is an alternative technique for separating ionic and strongly hydrophilic compounds without non-volatile ion-pair reagents, thus it is compatible with ESI/MS detection. The separation mode is similar to the normal-phase liquid chromatography (NP-LC) because a polar stationary phase is used but the strong organic mobile phase (typically acetonitrile, ACN) is replaced with an aqueous–organic mixture containing at least 5 % water.[190] The stronger solvent in the mobile phase is water, which promotes hydrophilic interactions between the analyte and a water-enriched hydrophilic stationary phase. The retention mechanism of HILIC is primarily described as a hydrophilic partition mechanism, in practice this "pure HILIC" mechanism is rare, as most of the HILIC separations are multimodal. In addition to the pure hydrophilic partition mechanism, the analytes interact with the stationary phase by ion-ion and hydrogen bonding

interactions.[191] Thus, retention can be influenced by the choice of the stationary phase. The same polar stationary phases as in the NP-LC are applicable in HILIC, such as bare silica and chemically bonded amino-, cyano-, diol-, amide- and switterionic- etc. phases.[190, 191]

The HILIC technique has many advantages compared to NP-LC and reversed-phase liquid chromatography (RP-LC), mostly due to the organic mobile phase. In contrast to NP-LC, the mobile phase includes water, but less than in the RP-LC. The elution order is reversed in the HILIC compared with RP-LC and the retention of polar compounds is better and non-polar compounds are retained poorly. MS response is better in HILIC compared with RP-LC because surface tension is smaller, therefore, the mobile phase is more volatile in ionisation.[189] High organic solvent proportions are used in HILIC mobile phases and volatile buffers lower surface tension, which increases ionisation efficiency in ESI.[189] Buffer salts such as ammonium formate or acetate[192] are commonly used to improve peak shape or retention in the HILIC separations, however, only low concentrations are suitable in ESI, high concentrations hamper the ionisation.[189] A current is proportional to flow rate in the ESI process. When current is increased by high flow rate, ion intensity decreases as larger droplets are formed which are less efficient for forming ions in the gas phase. Therefore a low flow rate is required.[189] The HILIC methods are generally coupled with ESI/MS detection for the analysis of highly polar and ionic compounds[192] such as nucleosides, nucleotides, carbohydrates, amino acids, peptides and proteins[191] including nucleotide phosphates[193] and nucleotide sugars.[194] For example, Preinerstorfer[194] and co-workers separated numerous hydrophilic intracellular metabolites including UMP, UDP and UDP-Gal with zwitterionic HILIC column and the metabolites were detected by LC-ESI tandem mass spectrometry (MS/MS).



Scheme 3.4. LgtC catalysed reaction when Rgal used as an acceptor

The aim was to develop HILIC method for the separation all reaction components of the LgtC catalysed reaction (Scheme 3.4), strongly hydrophilic UDP and UDP-Gal, less polar Rgal and the unknown product in single chromatographic run which is compatible with ESI/MS. Similar hydrophobicity (LogP) of UDP and UDP-Gal (Table 3.5) can make the separation of these compounds challenging. Additionally, more hydrophobic Rgal retains with a non-polar conditions compared to polar UDP and UDP-Gal.

Table 3.5. cLogP values of the analytes. Determined by ACD/3D Viewer Freeware 12.01

Analyte	Rgal	Rgal-gal[a]	UMP	UDP	UDP-Gal
LogP	-1.31	-2.41	-1.58	-4.27	-4.51

[a]Expected glycosylated product

The development of the chromatographic method was performed with a HPLC coupled with diode array detector (DAD). After the suitable separation was achieved, the HPLC method was transferred to HPLC coupled with ESI-MS/MS which was used for the analysis of the unknown product. The LCMS method development contained following steps:

- Investigation of the best chromatographic method for the separation of UDP-Gal, UDP and Rgal with bare silica, zwitterionic and amide modified silica HILIC columns by HPLC/UV
- Optimisation of the LgtC catalysed reaction conditions for the LC-MS run by HPLC/UV
- Optimisation of the LC-MS method for the analysis on the unknown reaction product
- Analysis of the reaction products by LC-MS/MS

3.3.4.1 Underivatised HILIC Silica bonded column

Bare silica or underivatised silica is the most popular packing material and originally used in NP-LC.[195] The surface of the silica is covered with hydrophobic siloxane (Si-O-Si) and different silanol groups (Figure 3.9) that can have different reaction and absorption activities.[195] Polar silanols interact with water molecules by hydrogen bonding from the mobile phase which leads to the formation of aqueous layer on the surface of the sorbent. The retention process is mostly hydrophilic interactions between the analytes and the silica stationary phase but different retention modes such as ion-exchange and some hydrophobic retention by the siloxane can occur depending on pH and salt concentration of the mobile phase.[196]

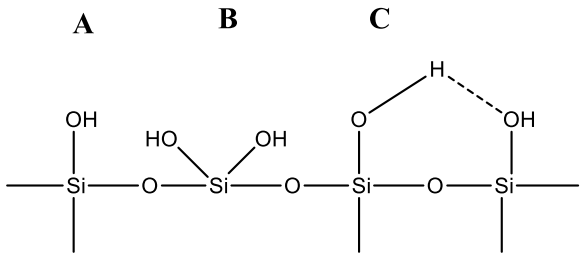


Figure 3.9. Different types of silanol groups on the silica surface: A) single, B) geminal and C) vicinal

Volatile mobile phases increase ionisation of analytes, therefore, silica packed columns are suitable in LC-MS systems.[196] Additionally, column bleeding is possible with derivatised silica packed columns[191] when functional groups can be released from the surface causing increased background noise on MS-based detectors but this cannot happen on bare silica columns. A disadvantage of bare silica columns is their stability to basic conditions, above pH 8, dissolution of silica skeleton can cause loss of column efficiency.[197] Bare silica HILIC columns are mostly used for the separation of various small polar compounds such as nucleosides and nucleic acid bases.[191, 196] Although, even there was a lack of applications for separation nucleotide sugars or nucleotide phosphates, bare silica bonded column was examined for the separation of UDP, UDP-Gal and Rgal.

The compatibility of Waters Atlantis HILIC silica column (3 µm particle size, 50 x 2.1 mm) was investigated by analysing a mixture of UDP-Gal, UDP and Rgal. The most hydrophobic Rgal was expected to elute first and then more hydrophilic UDP and UDP-Gal. Possibility to change composition of water during the run, initially the lowest water concentration was chosen. The detection of the components was started at 100 % of ACN:5 mM ammonium acetate (95:5) at pH 5.0.

The retention order was unexpected because in the HILIC systems, the retention order of polar analytes are usually opposite than in RP-LC separations.[192] UDP and UDP-Gal eluted as one peak in the column-void volume, approximately at one minute before Rgal at 2.5 min (Figure 3.10). The underivatised silica stationary phase is categorised as a weak hydrophilic phase,[198] therefore, strong adsorption or partition interactions between the hydrophilic UDP-Gal and UDP are prevented. Hydrophobic interactions between Rgal and the stationary phase, possibly explain the longer retention time.

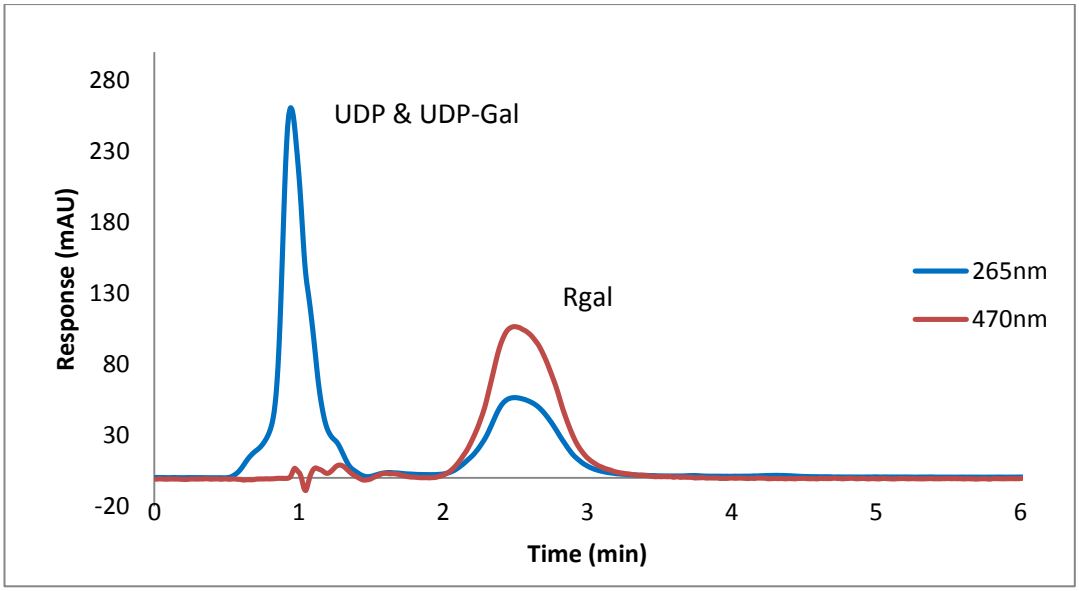


Figure 3.10. Separation of mixture of UDP, UDP-Gal and Rgal (100 µg/mL each) with bare silica column. UDP-Gal and UDP was detected at 265 nm (blue trace), and Rgal at 470 nm (red trace)

The retention was poor for all analytes even in the weakest mobile phase (least water). Therefore, the unmodified silica column was inappropriate for the analysis of the substances of the LgtC catalysed reaction. Based on the results, more hydrophilic stationary phase was required for the separation of the polar analytes.

3.3.4.2 ZIC-pHILIC column

Zwitterionic columns, commercially called ZIC-HILIC (silica gel support) and ZIC-*p*HILIC (polymer support) have been specifically developed for HILIC. A covalently attached sulfoalkylbetaine functional groups (Figure 3.11) on the column surface have a strong capability to adsorb water, thus this stationary phase is attractive for HILIC applications. The sulfoalkylbetaine ligand contains a negatively charged sulfonic acid group and a positively charged quaternary ammonium group, and the net charge is zero. Although, due to the large distance of sulfonic group from the surface of the silica than the ammonium group there is a low negative surface charge which is little affected by the pH value of the mobile phase. Therefore, ion-exchange is lower and electrostatic interactions are stronger compared to amide- and bare silica phases. ZIC-HILIC columns have been used to separate sugar phosphates,[199] nucleotide triphosphates[193] and nucleotide sugars,[194] therefore, the zwitterionic column was a promising alternative for analysis of substances of the LgtC reaction.

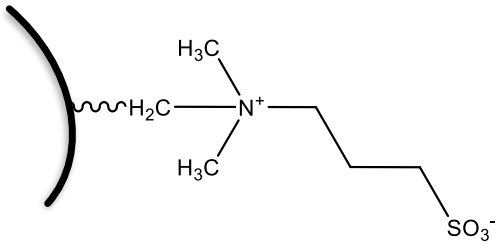


Figure 3.11. Sulfoalkylbetaine zwitterionic stationary phase

Polymeric-based SeQuant ZIC-*p*HILIC column (5µm particle size, 150 x 2.1mm) which is more hydrophilic than bare silica columns was examined. Two mobile phases were used: buffer A consisted of ACN and 10mM NH_4Ac (pH 4) (95:5, v/v) and buffer B consisted of ACN and 10mM NH_4Ac (pH 4) (50:50). The mixture of UDP, UDP-Gal and Rgal were run on the gradient of increasing concentration of aqueous content on the mobile phase (buffer B).

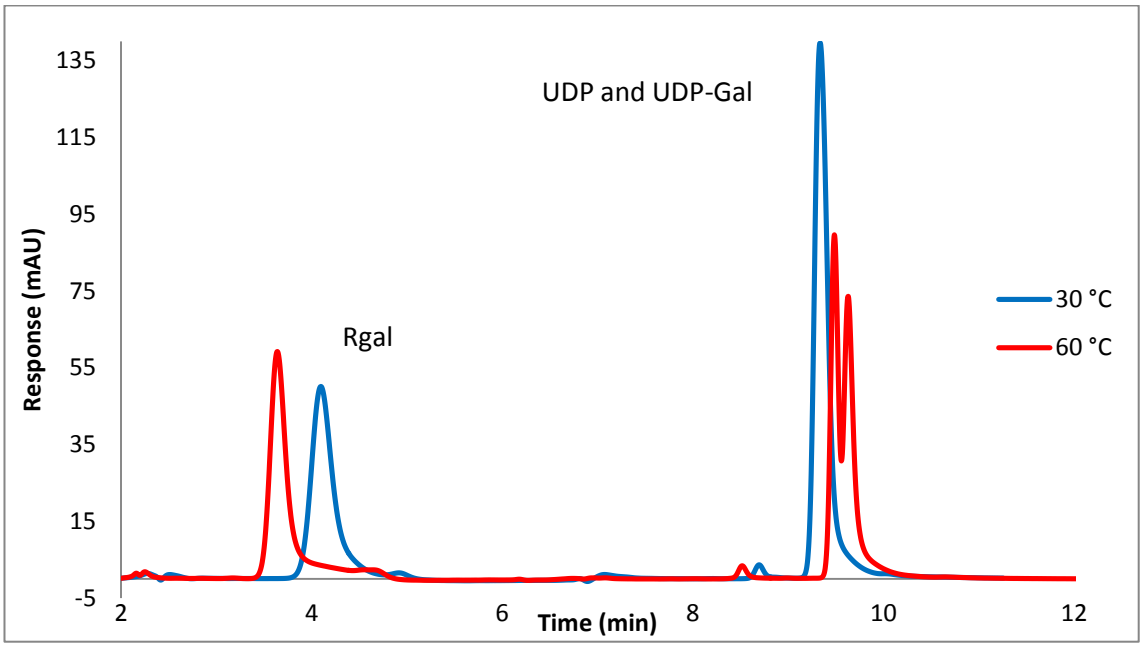


Figure 3.12. Mixture of UDP, UDP-Gal and Rgal (50 µg/mL each) analysed with ZIC-*p*HILIC column at 30 °C (blue trace) and 60 °C (red trace), detection wavelength 265 nm

Initially, the experiment was run at a column temperature of 30 °C and all the analytes were well retained, Rgal at 4 minutes and UDP/UDP-Gal at 9.3 minutes (Figure 3.12). Based on similar hydrophilicity of UDP-Gal and UDP (cLogP -4.51 and -4.27, respectively), it was assumed that the retention time was identical. Johnsen and co-workers discovered that the retention times of nucleotide triphosphates decreased with lower column temperatures and

they proposed the effect was due the dissociation of analyte/water clusters.[193] Physical properties of UDP are similar to the nucleotide triphosphates (*i.e.* UTP, nucleotide triphosphate) that were used in the Johnsen study. Therefore, it was assumed that by increasing column temperature the retention of UDP and UDP-Gal could change. Higher column temperature (60 °C) was examined and the results showed that the retention time of the peak at 9.3 minutes increased and the peak was split (Figure 3.12). In addition, the retention time of less hydrophilic Rgal reduced from 4 minutes to 3.6 minutes. The tall peak splitting at 60 °C demonstrated that UDP and UDP-Gal retained at 30 °C as one peak. The separation mechanism of UDP and UDP-Gal at elevated temperature is unclear. UDP-Gal is slightly more hydrophilic than UDP, however, UDP-Gal retained before UDP. Possibly ionic UDP can form hydrogen bonds with the aqueous layer and thus elute later. Or based on Johnsen's research, UDP/water clusters could be formed in higher affinity compared to UDP-Gal/water clusters and UDP is shielded from the mobile phase.[193]

Table 3.6. Measured parameters of UDP, UDP-Gal and Rgal (50 µg/mL each) analysed with various gradients at 60 °C. Resolution (R), tailing factor (T) and retention time (RT)

Gradient[a]	R, between UDP-Gal and UDP	T, Rgal	T, UDP	RT (min), Res-gal	RT (min), UDP-Gal	RT (min), UDP
1	1.16	0.81	0.46	4.53	20.73	21.20
2	0.87	0.82	0.45	4.55	20.18	20.62
3	0.54	0.69	poor	2.04	11.71	12.21
4[b]	0.98	0.83	0.49	4.51	19.67	20.06
5[b, c]	0.71	0.77	0.34	5.39	19.43	19.76

[a]Gradiends described in experimental chapter 3.6.2.3, [b]same gradient, [c]column temperature 30˚C

The separation of UDP and UDP-Gal was examined by running various gradient runs (presented in experimental chapter 3.6.2.3). More hydrophobic Rgal was retained in low aqueous content compared to hydrophilic UDP and UDP-Gal, therefore, gradients were set up by starting an isocratic 100 % buffer A for the first 5 minutes (gradient 4 & 5 in Table 3.6) or, the first 10 minutes (gradient 1). The retention time of Rgal with 100 % buffer A was 4.5 minutes and remained the same when water content was slowly increased (gradient 2). When the water content in the mobile phase was increased rapidly (gradient 3), all analytes had a shorter retention times. However, different aqueous content did not significantly improve the resolution between UDP and UDP-Gal (Table 3.6). The best resolution was achieved with high aqueous content at 25 % buffer A (Table 3.6, gradient 1 & 4). The back

pressure significantly limited the aqueous content on the gradient run and the maximum buffer B used was 75 %. Therefore, the ZIC-*p*HILIC column did not meet the satisfactory resolution (R >1.5) between UDP and UDP-Gal.

3.3.4.3 Amide bonded HILIC column

The ZIC-*p*HILIC column was unsuitable for the separation of UDP and UDP-Gal because of too high back pressure in high water content, therefore, the last option was to investigate an amide modified silica HILIC column. Amide HILIC columns have been mainly utilised for analysis of peptides and monosaccharides,[191] although also nucleotides[200] and nucleotide sugars[201, 202] have been successfully separated. The stationary phase contains a covalently bonded amide (e.g. carbamoyl) groups via short aliphatic linkage to the silica backbone (Figure 3.13). Compared to amino silica columns, the amide groups are less reactive and the retention of the ionic analytes is less prone for ion interactions.[203]

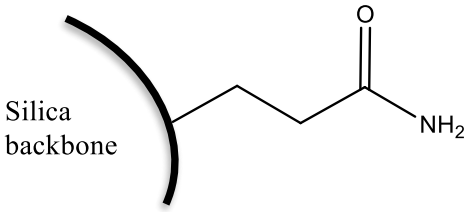


Figure 3.13. Amide modified silica stationary phase

Waters ACQUITY BEH Amide column (particle size 1.7µm, 150 x 2.1 mm) which is designed for ultra-performance liquid chromatography was examined. Mixture of UDP, UDP-Gal and Rgal was prepared in water:ACN (20:80) and analysed on a gradient run of increasing water content. Buffer A consisted of ACN: 10mM NH_4Ac (pH 9) (95:5) and buffer B consisted of ACN: 10mM NH_4Ac (pH 9) (50:50). Two sharp peaks at 3.4 and 17.8 minutes were observed (Figure 3.14 A). ZIC-*p*HILIC and amide column columns have similar retention behaviour,[203] therefore, the retention order of UDP and UDP-Gal was expected to be identical, and the broad peak at 18.7 minutes was assumed to be UDP.

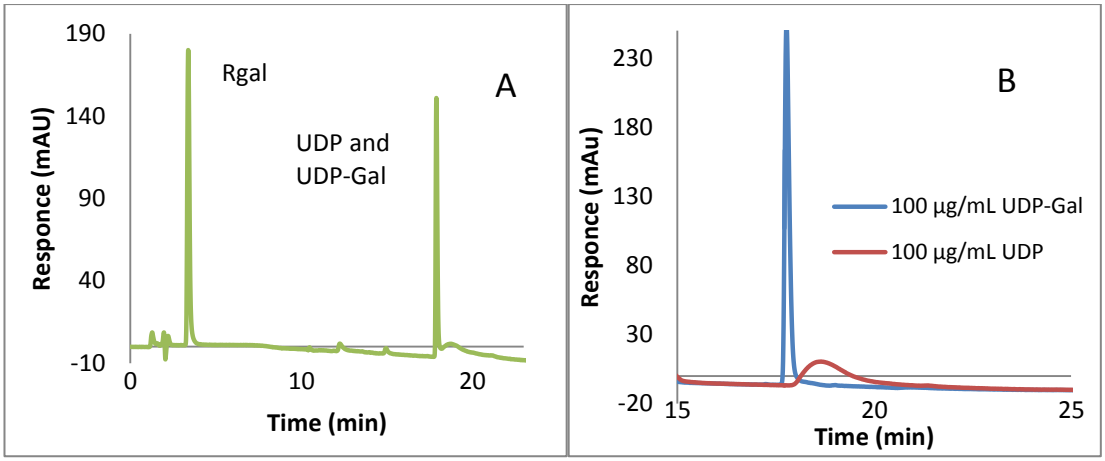


Figure 3.14. A) Mixture of Rgal, UDP-Gal and UDP (green trace), 50 µg/mL each B) UDP (red trace) and UDP-Gal (blue trace), 100 µg/mL each examined with two separate HPLC runs. Detection wavelength 265 nm

The broad peak was identified as UDP (Figure 3.14 B) by analysing UDP and UDP-Gal (100 µg/mL each) at identical conditions as the mixture with two separate chromatographic runs. Peak broadening can be caused by a stronger sample solvent than the mobile phase solvent.[204] In this case, the stronger solvent was water and UDP was diluted in water:ACN (20:80). The aqueous content was approximately 55 % at the retention time of UDP. The aqueous content of buffer B was then increased from 50 % to 70 % to investigate the influence of stronger solvent on peak shape.

The peak shape of UDP improved significantly and the response of both UDP and UDP-Gal peaks increased (Figure 3.15, A & B) with higher aqueous content in the mobile phase. The retention time decreased with increasing polarity from 17.8 to 16.3 minutes, and from 18.7 to 16.7 minutes for UDP-Gal and UDP, respectively. The peak shape and sensitivity of Rgal was maintained in higher aqueous content and surprisingly, the retention time increased from 3.5 to 6.1 minutes (Figure 3.15, B & C). The retention of Rgal is reversed-phase-like, as elution was shorter in a more hydrophobic environment. The peak in front of UDP-Gal was UMP, the hydrolysis product of UDP-Gal. Retention and separation of Rgal, UDP-Gal and UDP were satisfactory and the retention time of the product of the LgtC catalysed reaction was unknown at this stage. The retention time of the product was expected to be between Rgal and UDP-Gal based on the hydrophilicity. Therefore, the gradient was extended from 15 to 20 minutes (Figure 3.15, C) and as a result UMP, UDP-Gal and UDP retained later, at 18.2, 18.7 and 19.2 minutes, respectively. Additionally, resolution between UDP and UDP-Gal increased from 2.25 to 2.75.

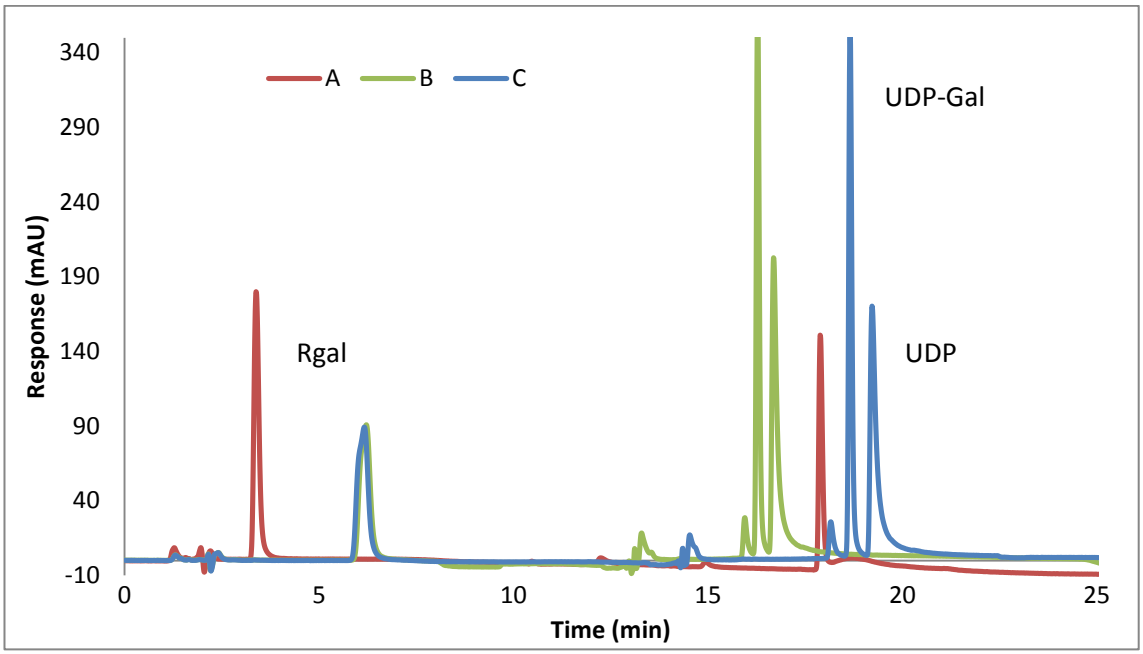


Figure 3.15. Separation of the mixture of Rgal, UDP and UDP-Gal (50 µg/mL each). Overlay of HPLC runs in A) gradient 0-100 % of 10 mM NH_4Ac:ACN (50:50) over 15 minutes against buffer A (10 mM NH_4Ac:ACN (5:95)) B) gradient 0-100 % of 10 mM NH_4Ac:ACN (70:30) over 15 minutes against buffer A and C) the same as trace B but gradient over 20 minutes. Detection wavelength 265 nm

It has been shown that sample diluent has a significant influence on peak shape. Ruta and co-workers studied peak shape in various sample diluents systematically with HILIC mode and they recommended to dilute samples in the lowest aqueous content possible to maintain suitable peak shape.[204] The peak shape of UDP was tailing and Rgal was slightly split, thus various water contents in the sample diluents were studied to improve the peak shape. Mixtures of Rgal, UDP and UDP-Gal were prepared in 10 %, 20 % and 30 % water in ACN.

The most symmetrical Rgal peak shape was obtained at 10 % of water and peak symmetry was lost with increasing water content (Figure 3.16 A) resulting in a split peak at 30 % water. Rgal retained at approximately 20 % of aqueous content in the mobile phase. The poor peak shape was possibly affected by the difference in the solvent strength between the mobile phase and the sample solvent. The opposite was observed with UDP-Gal and UDP (Figure 3.16 B). The response increased with increasing water content but remained the same at 20 % and 30 %. The recommendation of Ruta's research "the lowest water content in dissolution solvent maintain suitable peak shapes" [204] was not suitable for the analysis of strongly polar UDP-Gal and UDP. The possible explanation is the difference between the low water proportion in the sample dissolution solvent and the high water content (70 %) of the mobile phase at the

retention time of UDP-Gal and UDP. The aqueous layer of the stationary phase may be disturbed by the water content in the dissolution solvent which have ability to form hydrogen bonds and ionic UDP may be affected. Subsequent experiments were decided to be performed at 20% water because the response and the peak shapes of all analytes were agreeable.

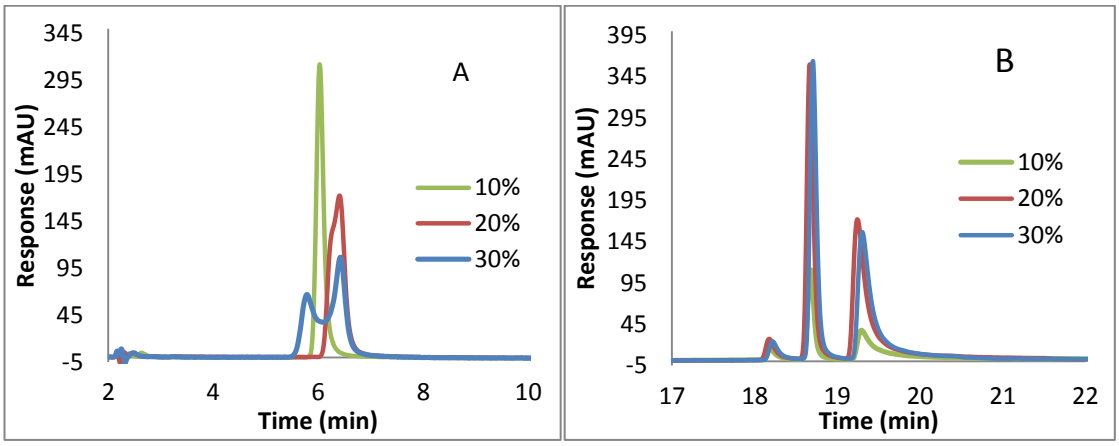


Figure 3.16. Influence of various water content in the sample solvent, 10 %, 20 % and 30 % of water in green, red and blue traces, respectively. A) with Rgal, detected at 470 nm and B) with UMP, UDP-Gal and UDP (respectively, in elution order) were detected at 265 nm

In order to identify the unknown product of the LgtC catalysed reaction when Rgal was used as an acceptor from the HPLC trace, a test of the precision of peak retention of Rgal, UMP, UDP-Gal and UDP was performed. In addition, resolution between UMP and UDP-Gal, and between UDP-Gal and UDP were investigated. Six replicate injections from homogenic sample with fixed HPLC conditions were performed.

Table 3.7. Reproducibility of retention time (RT) and resolution (R) for Rgal, UMP, UDP-Gal and UDP (50 µg/mL each), (*n*=6). Experiment performed with HPLC gradient: 0-100 % of buffer B (10 mM NH_4Ac:ACN (70:30)) over 20 minutes against buffer A (10 mM NH_4Ac:ACN (5:95))

	RT (min) Rgal	RT (min) UMP	RT (min) UDP-Gal	RT (min) UDP	R between UMP and UDP-Gal	R between UDP-Gal and UDP
Average	6.20	16.71	17.42	17.86	2.34	2.40
RSD%	1.88	2.99	0.29	0.26	0.99	2.04

The reproducibility of retention time was satisfactory for UDP-Gal and UDP (Table 3.7), however, the precision was not as repeatable for Rgal and UMP. The resolution between UMP and UDP-Gal peaks, as well as between UDP-Gal and UDP peaks were excellent.

3.3.5 Optimisation of LgtC catalysed reaction conditions for LC-MS

The HPLC method with the amide modified silica HILIC column was developed for the analysis of Rgal, UDP-Gal and UDP. The unknown product of the LgtC catalysed reaction when Rgal was used as an acceptor, was eluted between UDP and Rgal with the IPC method (Figure 3.8). Based on the polarity of the product, it was assumed that the retention time would be between Rgal and UMP with the HILIC method. Before coupling the HPLC method with the LC-MS instrument it was essential to determine the retention time of the product and optimise the reaction conditions to ensure that an appropriate amount of the product was formed.

The reaction conditions were directly adapted from the glycosidase coupled assay (Chapter 3.3.3) and the reaction products were separated and analysed with the amide HILIC method. The total volume of the reaction was 200 µL and the reaction was stopped by adding 800 µL ACN, thus, the aqueous proportion of the sample was 20 %. The control was prepared in the absence of donor (UDP-Gal) and treated as the normal reaction mixture.

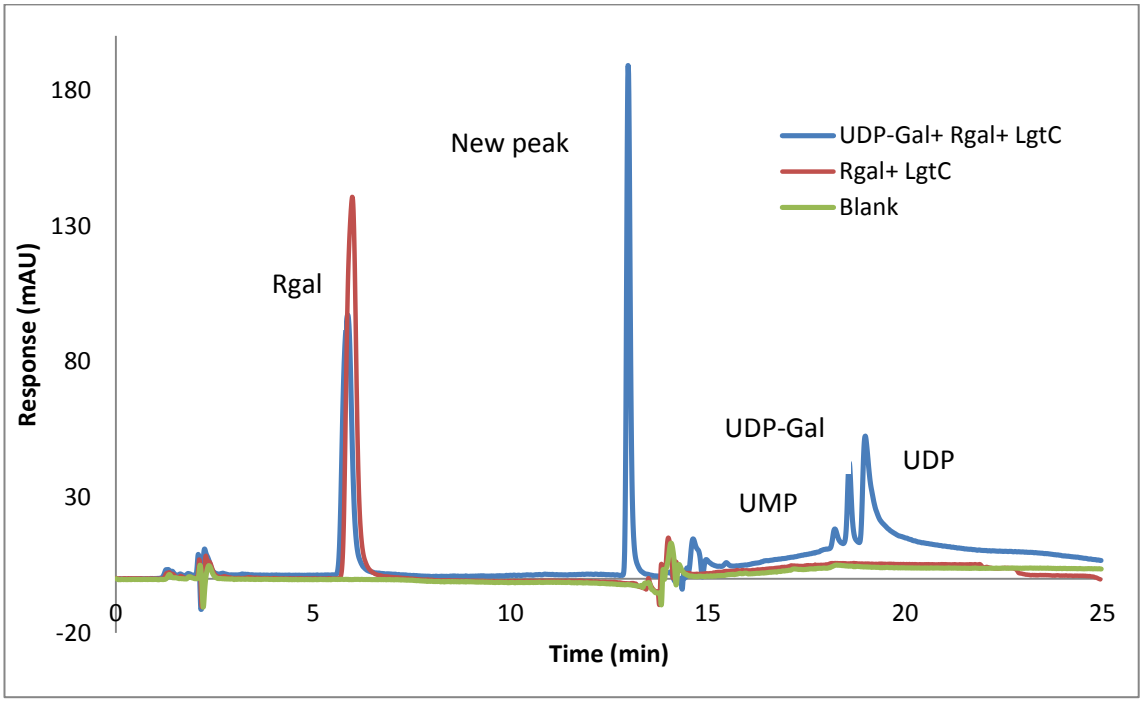


Figure 3.17. HPLC chromatogram of LgtC catalysed reaction, Rgal as an acceptor. A new peak appeared at 13 minutes (blue trace), blank and control traces in green and red, respectively. Compounds separated with Amide HILIC column and detected at 265 nm

Rgal peak was decreased in the reaction compared with the control (no UDP-Gal) and a new peak retained at 13 minutes (Figure 3.17). Based on the retention time, the new peak is likely to be the glycosylated Rgal as the expected product is more hydrophilic than Rgal. The hydrolysis product of UDP-Gal, UMP was formed during the reaction and the UMP peak was observed before UDP-Gal. The baseline was noisy from 13.3 to 14.5 minutes, however, the irregularity was caused by the gradient because the same was observed in the blank. However, the retention time does not reveal how many galactose moieties are transferred and the product peak required identification. Therefore, these HPLC and reaction conditions were adapted to the LC-MS system.

3.3.6 Analysis of the reaction product by LC-ESI-MS/MS

The developed HILIC method was directly suitable for ESI because the volatile mobile phase and a low flow rate were used. The amide modified silica HILIC column was installed to HPLC-ESI-MS/MS instrument which can perform two stage mass analysis to examine selectively the fragmentation of particular ions from the mixture.[205] The LC-MS ionisation conditions were optimised prior to the analysis of the reaction product of the LgtC catalysed reaction.

Fraction patterns of the standard solutions of UMP, UDP, UDP-Gal and Rgal were examined and the optimum product ion was chosen for each analytes to give the strongest signal by varying the collision energy (CEn). UMP, UDP and UDP-Gal were ionised with the negative mode and Rgal with the positive mode. UMP was determined by detecting the cleavage of the nucleotide base, UDP detected by the cleavage of single phosphate group, UDP-Gal detected by the cleavage of sugar phosphate and Rgal by the sugar cleavage (Table 3.8). Standards were not available for the expected glycosylated product, however the product was assumed to be ionised in positive ionisation mode, the same as for Rgal.

Table 3.8. Optimised LC-ESI-MS/MS parameters, fraction patterns, molecular weight (Mw), collision energy (CEn) and retention times (RT) of the standard solution of the analytes

Compound	Mw	Parent ion	Product ion	Ionisation Mode	CEn	RT
Rgal	375.3	375.8	214.1	Pos	25	4.26
UMP	324.2	323.3	211.1	Neg	28	12.27
UDP-Gal	566.3	565.2	323.0	Neg	25	12.61
UDP	404.2	403.2	305.1	Neg	30	13.84

LC-MS run of the standards did not show clear peaks on the total-ion spectrum because MS detection was not optimised. However, MS/MS analysis of the standards illustrated clearly visible peaks (Figure 3.18). The retention order of the standards was the same as with the standard HPLC time, however, the retention times were reduced which was possibly caused by the different LC instrument used (pump, tubing etc.).

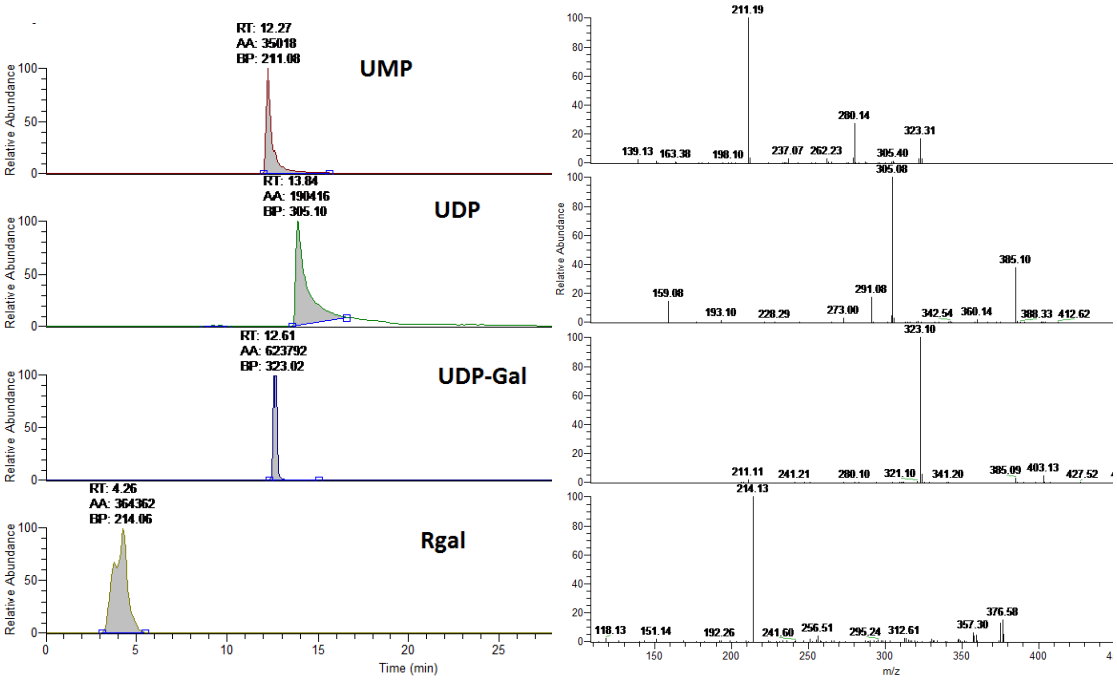


Figure 3.18. LC-ESI-MS/MS analysis of the standards of UMP, UDP, UDP-Gal and Rgal. The *m/z* fragmentation of the analysed peaks presented on next to the peak for each standard

The aim was then to identify the molecular weight of the product of the LgtC catalysed reaction with LC-MS/MS. All the substances of the reaction were separated with the previously described method (Chapter 3.3.5). The retention time of the unknown product was assumed to be between Rgal and UMP. Based on the similar physical properties of the expected unknown product (Rgal-gal) and the acceptor Rgal, the ionisation was assumed to occur in positive mode. The product peak was identified by entering the molecular weight of the expected product manually.

Initially it was investigated if the unknown product may be Rgal-gal. Based on a molecular weight of Rgal-gal, a peak eluted at 8.1 min (Table 3.9 and Figure 3.19) which was ionised with the positive mode. The observed *m/z* of 538.3 was assumed to be the parent ion and a corresponding product ion was formed by cleavage of digalactose, *m/z* of 215.2. The experiment was repeated and by comparing *m/z* with the controls: UDP-Gal and Rgal in the

absence of LgtC and UDP-Gal or Rgal in the presence of LgtC. The results confirmed that the observed peak at 8.1 minutes was not an artifact because the *m/z* was not found in the controls. However, because the expected product was Rgal-gal, the results did not explain the high consumption of UDP-Gal in the reaction. It was suggested that Rgal-gal could serve as an acceptor for LgtC and form multiple products by elongating the galactose chain. Thus, all possible glycosylated products up to four transferred galactose to Rgal with corresponding masses were examined with LC-MS/MS. One new peak at lower intensity was retained at 10.2 minutes, and the parent and the product ions, 700.0 and 214.8, respectively, suggested that the peak is Rgal-gal-gal (Table 3.9 and Figure 3.19).

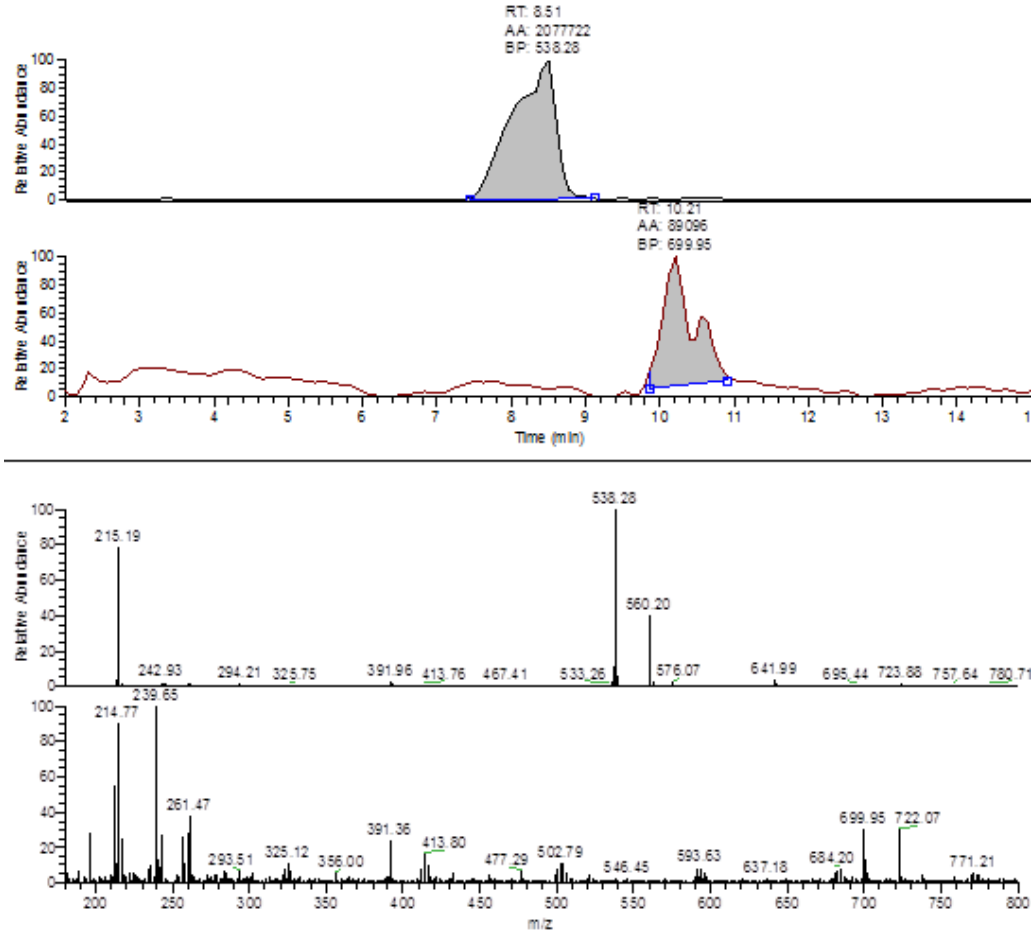


Figure 3.19. Glycosylated products of LgtC catalysed reaction: Rgal-gal (8.5 minutes) and Rgal-gal-gal (10.2 minutes), and the corresponding masses below, respectively

The LC-ESI-MS/MS results suggested that two products were formed in the LgtC catalysed reaction when Rgal was used as an acceptor. The mass of the products corresponded to Rgal-gal and Rgal-gal-gal. In addition, the products retained with increasing hydrophilicity order which confirms the observed results and Rgal-gal-gal was detected with a lower intensity

than Rgal-gal-gal. The MS/MS is more sensitive than the UV-detection, and thus only one product peak was observed during HPLC analysis.

Table 3.9. Identified products of the glycosylated product of the LgtC catalysed reaction when Rgal was used as an acceptor

Compound	Mw	Parent ion	Product ion	Mode	RT
Rgal-gal	537.5	538.3	215.2	Pos	8.1
Rgal-gal-gal	699.6	700.0	214.8	Pos	10.2

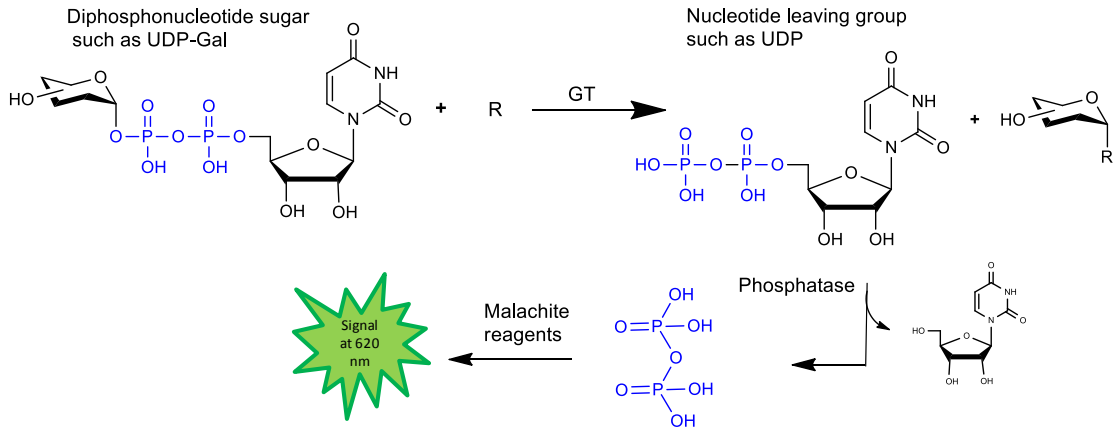
The observed results suggested that LgtC utilises both, Rgal and Rgal-gal as acceptors and two galactosylated products were detected by LC-MS/MS. Antoine and co-workers found multiple side products during the enzymatic synthesis of globotetraose (GalNAc-β1,3-Gal-α1,4-Gal-β1,4-Glc).[206] The findings suggested that LgtC used globotriose (Gal-α1,4-Gal-β1,4-Glc) as an acceptor after complete exhaustion of lactose resulting in polygalactosylated compounds. Tetra, penta and hexa-hexose derivatives of globotetraose were identified.[206] The glycosidase coupled assay was performed in the absence of lactose, therefore, Antoine's results suggested that Rgal-gal could serve as an acceptor for LgtC. Multiple polygalactosylated products were possibly formed causing a rapid exhaustion of UDP-Gal, therefore, the reaction completion was not observed with the glycosidase coupled assay or HPLC-based assay when Rgal served as an acceptor. In addition, during the optimising the LgtC reaction conditions for the LC-MS experiments, UMP, the hydrolysis product of UDP-Gal was formed in the assay conditions (Figure 3.17). The glycosidase coupled assay based on the detection of the non-glycosylated acceptor and in the assay format, there was no method for detecting the hydrolysis of the donor substrate, UDP-Gal, which could have subtracted from the fluorescence signal. Therefore, the hydrolysis of UDP-Gal which has an impact on the depletion of UDP-Gal is impossible to quantify. Due to these difficulties, the development of the glycosidase coupled assay was brought to an end at this stage.

3.4 Phosphatase coupled biochemical assay

After the complexities encountered in the development of glycosidase coupled assay, an alternative biochemical assay was still required for the investigation of potential GT inhibitors. During the course of this project, a novel phosphatase coupled assay was described and this assay was optimised in house with β1,4-GalT and LgtC.

3.4.1 Principle of the assay

Wu and co-workers developed an universal phosphatase coupled assay (PCA) that was applied to several GTs including β1,4-GalT.[85] The assay is based on specific phosphatase that releases an inorganic phosphate from the leaving groups of GT reactions and the released phosphate group is quantitatively detected by colorimetric malachite reagents (Scheme 3.5). Malachite green is a widely used colorimetric reagent that allows detection of inorganic phosphate due to the formation of a phosphomolybdate malachite green complex with absorption maximum at 620 nm.[207, 208]



Scheme 3.5. Phosphatase coupled GT assay can be applied to any GT reaction where the leaving group contain a removable phosphate (*i.e.* UDP from UDP-Gal). (The scheme is modified from reference[85])

In this discontinuous assay, the enzymatic reactions are stopped after desired time by adding highly acidic malachite reagents which denature the protein. The PCA is suitable for Leloir GTs that generate nucleotides as the leaving groups such as UDP,[5] therefore, the assay is adaptable for LgtC. The assay is an alternative choice for compound characterisation after

HPLC because the PCA protocol is rapid, simple to perform and the assay can be performed on 96 well microplate that allows determination of multiple compounds simultaneously.

The PCA optimisation with β1,4-GalT was performed by Dr Evitt (the assay optimisation is described in appendix 1). Dr Evitt discovered an optimal phosphatase that hydrolysed inorganic phosphate specifically from UDP, and a carrier protein that improved reproducibility of the assay performance. The phosphatase coupled assay was suitable for detecting the secondary product (UDP) of LgtC catalysed reaction and no suitable assay existed to study activities of LgtC. Therefore, the aim was to apply the improved PCA to LgtC for the first time. The following parameters were investigated to make the PCA suitable for the investigation of enzymatic kinetic parameters and for inhibitor characterisation with LgtC:

- The acceptor needed to be changed
- Influence of protein carrier and Mn^{2+}
- Influence of DMSO, Mn^{2+} and TX-100 on enzymatic kinetic parameters

3.4.2 Optimising the assay conditions with LgtC

The original phosphate assay was performed in the absence of a carrier protein.[85] During the assay optimisation a carrier protein (hen egg-cell lysozyme, CEL) was discovered which significantly improved the assay performance with β1,4-GalT (appendix 1). The assay optimisation with LgtC was initiated by an experiment with increasing concentrations of LgtC in the presence and in the absence of the carried protein to study if LgtC was dependent on the carrier protein. In addition, the acceptor of β1,4-GalT, GlcNAc, was changed to lactose at the same concentration that was used in the HPLC-based method.

The results indicated that the activity of LgtC improved significantly in the presence of CEL (Figure 3.20). The observed results indicated that the carrier protein possibly protected LgtC from attaching to plastic surfaces of the microplate and the pipette tips. Therefore, subsequent experiments were performed in the presence of CEL.

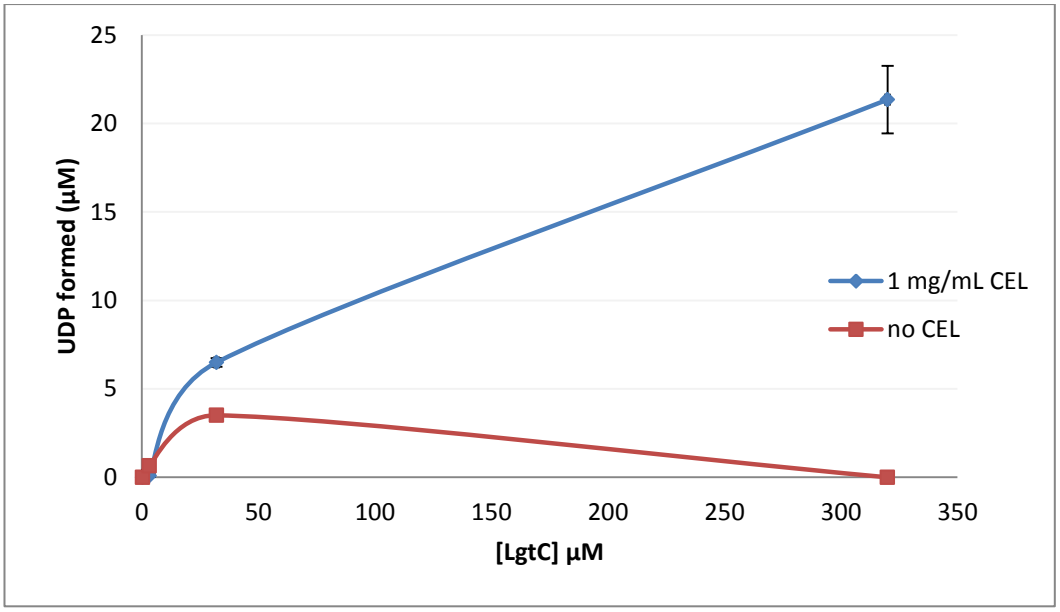


Figure 3.20. LgtC catalysed reaction in the presence and in the absence of CEL

The background hydrolysis of UDP-Gal and ~7 % impurity was found during the PCA development was subtracted from the main signal in the assay set up with β1,4-GalT (appendix 1). The background control included UDP-Gal and β1,4-GalT in desired concentrations in the absence of acceptor. However, lactose, the acceptor of LgtC contained inorganic phosphate as an impurity and the background control was unsuitable to perform in the absence of the acceptor because the signal of the control decreased significantly. Therefore, whether the background control could be performed in the absence of LgtC the hydrolysis of UDP-Gal was investigated in the presence and in the absence of LgtC (Figure 3.21).

The results indicated that hydrolysis of UDP-Gal remained the same in the absence and in the presence of LgtC during the incubation time of 20 minutes. Because LgtC was incapable of hydrolysing UDP-Gal the background hydrolysis was performed in the absence of LgtC instead. The results also indicated that DTT in LgtC solution did not prevent the colour development at this concentration range because the absorbance signal obtained was similar in the presence and in the absence of LgtC (Figure 3.21).

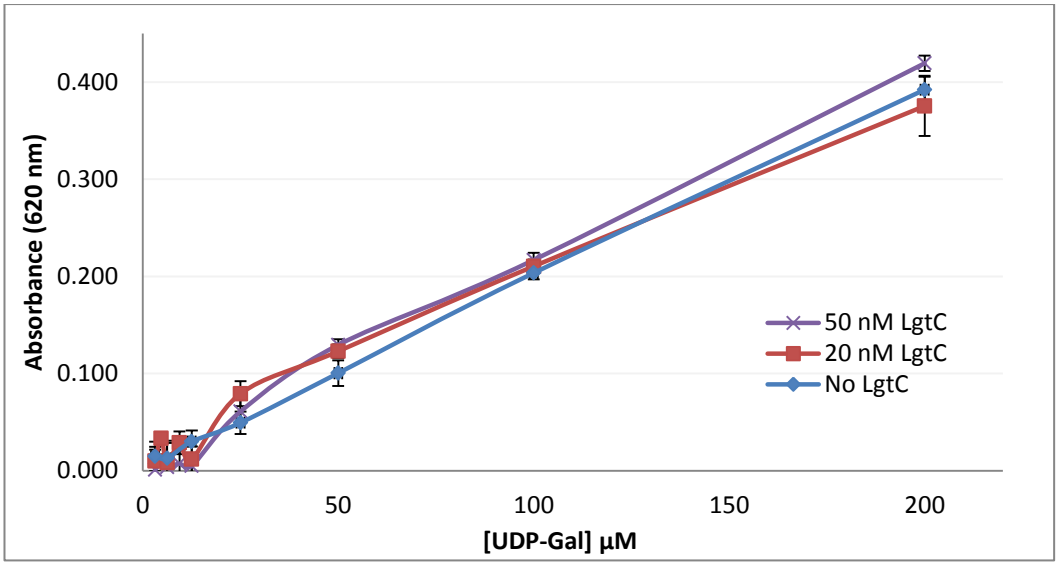


Figure 3.21. UDP-Gal hydrolysis in the presence and in the absence of LgtC at 20 nM and 50 nM

The influence of manganese concentration on the background hydrolysis and activity of LgtC was investigated. All assay parameters were fixed, except manganese concentration. As demonstrated in the Figure 3.22, the background hydrolysis was independent of manganese concentration and activity of LgtC was dependent on manganese concentration. The maximum activity of LgtC increased significantly in the presence of $MnCl_2$ and maximum activity was observed at 5 mM of $MnCl_2$. Moreover, the results showed LgtC is still active without the addition of $MnCl_2$ due to the manganese present in the LgtC stock solution (chapter 2.7.1).

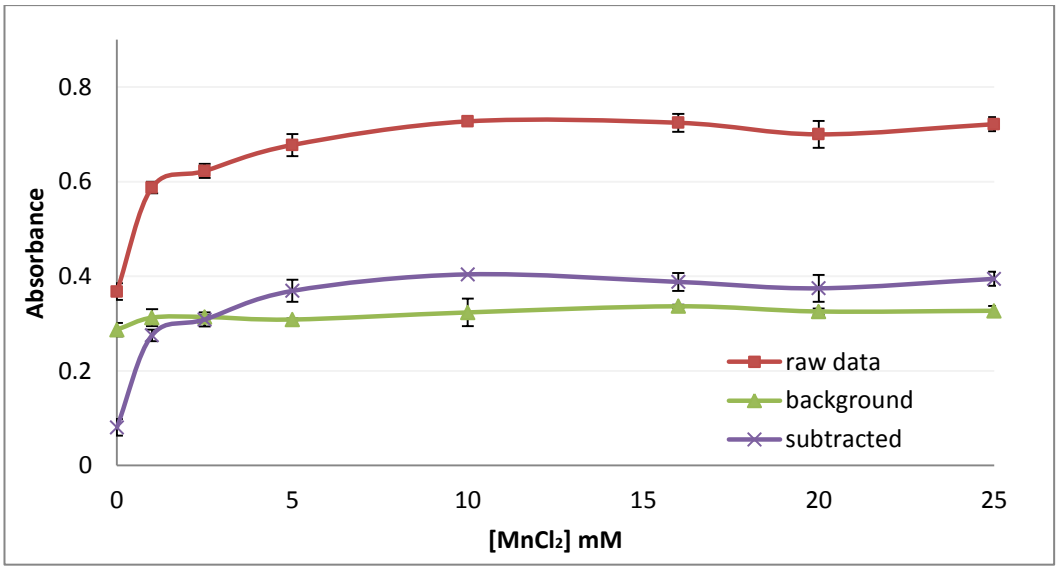


Figure 3.22. Formation of phosphate with increasing $MnCl_2$ concentration and at fixed concentrations of UDP-Gal and LgtC

The PCA was designed for use in the inhibitor characterisation experiments. Many small molecular compounds are prone to inhibit enzymatic activity by sequestering proteins which can cause false positive results. Therefore, influence of TX-100, which has been proved to prevent compound aggregation,[209] was investigated with the PCA. The colour development of malachite reagents was examined at various UDP and TX-100 concentrations (Figure 3.23). The results illustrated that TX-100 did not affect the colour development in the measured range. The same amount of phosphate was formed in the presence and in the absence of TX-100.

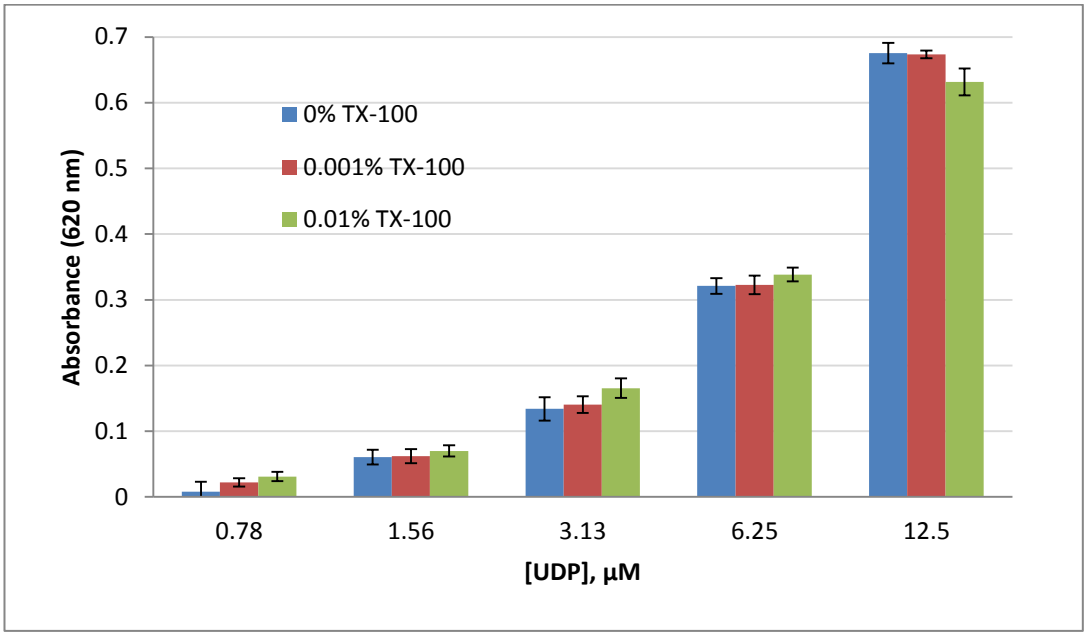


Figure 3.23. Colour development of the malachite reagents at fixed phosphatase concentration and in various TX-100 and UDP concentrations

In addition, influence of TX-100, DMSO and without addition of $MnCl_2$ to the enzyme kinetic parameters such as K_m, V_{max} and K_{cat} were investigated. The concentration of LgtC was selected where the UDP-Gal conversion to the product was 10 % or less. In the standard assay conditions K_m was 11.8 µM, that is comparable with the literature value, 18 µM.[16] The effect of DMSO to K_m value was determined at 10 % (v/v) and the K_m value decreased from 11.8 µM to 3.5 µM indicating that DMSO activates LgtC. This effect may due to favourable conformation changes on the donor binding site in the presence of DMSO.[210] The K_m value remained the same in the absence and in the presence of 0.01 % (v/v) TX-100. Additionally, K_m was in the absence of $MnCl_2$ addition to evaluate if the PCA could be performed with similar conditions than the LDA (chapter 2). The result demonstrated that it was possible to determine K_m without $MnCl_2$ addition. However, the efficiency of LgtC was lower, and a

higher LgtC concentration was required to obtain the optimum turnover which increased the K_m value.

Table 3.10. Enzyme kinetic values of UDP-Gal in various assay conditions (n=3)

Conditions	K_m (μM)	V_{max} (μM/min)	K_{cat} (s^{-1})
Standard[a]	11.8 ± 2.7	0.43 ± 0.03	0.36 ± 0.02
+ DMSO	3.5 ± 1.9	0.38 ± 0.05	0.32 ± 0.05
+ DMSO, No Mn^{2+}	27.4 ± 4.1	0.34 ± 0.01	0.11 ± 0.01
+DMSO +TX-100	3.5 ± 1.7	0.45 ± 0.04	0.30 ± 0.11

[a]Standard assay conditions, reaction mixture in 13 mM HEPES, pH 7.0: Lactose (2 mM), phosphatase (10 U/mL), CEL (1.0 mg/mL), $MnCl_2$ (5.0 mM), UDP-Gal (0.79-200 μM) and desired LgtC concentration. DMSO concentration 10 % (v/v) and TX-100 concentration 0.01 % (v/v)

The optimisation experiments demonstrated that the PCA can be used for the characterisation of LgtC activities with similar conditions to those with β1,4-GalT. The PCA is simple to perform and due to the tolerance for DMSO and TX-100, it is suitable for compounds characterisation against LgtC. The PCA can be performed with the similar assay conditions to the LDA, including the same DMSO, TX-100 and $MnCl_2$ concentrations. An advantage of the PCA is the simple and rapid assay protocol. In addition, the background hydrolysis of UDP-Gal can be determined, and this is advantageous. However, a minor disadvantage of the assay is that the microplate holds many controls. In addition to background hydrolysis controls and phosphatase controls, the wells on the outer edge are left empty because the reaction velocity was lower, possibly due to different temperature.

3.5 Summary and conclusions

Three different assays have been investigated for the development of a biochemical assay for LgtC. After investigating the HPLC-based method and the novel glycosidase couple assay, the phosphatase couple assay was the option with the most potential to study activities of LgtC.

The HPLC-based method was successfully developed for the separation of UMP, UDP-Gal and UDP in a single chromatographic run with total analysis time of 22 minutes. The system suitability tests demonstrated a satisfactory system performance. The disadvantages appeared when the HPLC method was applied to monitor the LgtC catalysed reaction. The enzymatic reactions were stopped by placing the samples in dry ice and the most time consuming step was the sample thawing before each HPLC injection. The kinetic parameters,

such as K_m, require numerous data points, therefore, the step must be performed multiple times which leads to a long analysis time. Stopping the enzymatic reaction chemically in order to achieve more automated assay protocol was investigated. The reaction stopping study was carried out with ethylenediaminetetraacetic acid (EDTA) and with HCl to inactivate LgtC by chelating available Mn^{2+} or by denaturation, respectively. In both cases LgtC remained active. Another option for stopping the enzymatic reactions is by diluting the samples, however, this manner would further increase LOQ. The LOQ could be decreased by increasing the injection volume, however, this could have a negative impact on the peak shape. Therefore, the reactions were stopped by freezing.

The advantage of the HPLC-based method is that the enzymatic reactions can be monitored without additional derivations and the hydrolysis of UDP-Gal can be monitored simultaneously with the donor substrate and the reaction product. The usefulness of the HPLC-based method was proved during the development of the novel glycosidase coupled assay. The HPLC-based method provided useful information about the consumption of UDP-Gal, whereas the glycosidase coupled assay detected only the amount of unreacted acceptor.

The goal to develop a novel GT assay based on an unnatural fluorescent acceptor and coupling of the GT reaction to a glycosidase reaction was not achieved. During the assay development, unforeseen enzymological features of LgtC were discovered. LC-MS/MS was utilised for the investigations of unknown products and the results suggested that LgtC transfers one and two galactose moieties to the unnatural acceptor forming Rgal-gal and Rgal-gal-gal. Antoine and co-workers discovered similar behaviour with LgtC during the enzymatic synthesis of globotetraose.[206] They observed polygalactosylated compounds which suggested that globotriose serves as an acceptor of LgtC after complete exhaustion of lactose. Previous studies had also shown that D-galactose serves as an acceptor with much lower efficiency compared to lactose.[174] The glycosidase coupled assay was performed in the absence of lactose which suggests that α-linked galactose could serve as an acceptor for LgtC in the assay conditions. Additionally, resorufin-β-galactopyranoside (Rgal) is probably a less efficient acceptor than galactosylated product, therefore, Rgal remains unreacted. These results suggested that the assay design for LgtC is essentially performed in the presence of excess lactose that prevents the α1,4-linked galactose serving as an acceptor. To date, polygalactosylation has not been investigated systematically with LgtC. However, during the development of the glycosidase coupled assay, an assay based on the same principle was

reported. LgtC activities were examined by using an unnatural fluorescent acceptor, 4-nitrophenyl-β-D-galactopyranoside.[211] Questionably, a high K_m value of UDP-Gal (5.85 mM at 25 μM of acceptor and 5-100 μM of UDP-Gal) was reported which suggests that galactosylated 4-nitrophenyl-β-D-galactopyranoside may have served as an acceptor. However, the reaction products of LgtC were not further analysed in the report and it is unclear whether polygalactosylation was observed.

Three HILIC columns were examined for the separation of UDP, UDP-Gal and Rgal. Satisfactory separation was obtained with the amide-bonded column which was then utilised for monitoring the LgtC catalysed reaction on the LC-MS experiments. The amide column is an alternative for IPC methods, the nucleotide sugars and nucleotides were retained without ion-pair reagents and high buffer concentrations were not required. The IPC method was performed in high phosphate buffer concentration and, compared with the amide HILIC column, the separation of the nucleotide sugars can be performed in lower buffer concentration. This can increase the column lifetime, and additionally, possible buffer precipitation in the HPLC system can be minimised.

The novel phosphatase coupled assay was developed by Wu and co-workers[85] and after the assay was optimised with β1,4-GalT in house, it was applied to LgtC. The reaction monitoring based on a phosphatase that releases inorganic phosphate from UDP during GT reaction and the generated phosphate group is detected by phosphate detecting reagents. The assay conditions such as CIP, CEL, $MnCl_2$, DMSO concentrations, incubation time and incubation temperature, remained the same with LgtC as used with β1,4-Gal. The enzyme kinetic studies of LgtC were carried out on multiwell microplates and the results demonstrated that the PCA was robust and sensitive technique. Additionally, the assay protocol is simple to perform and it is significantly less time-consuming than the HPLC-based method. For example, K_m with LgtC can be performed with eight concentrations of UDP-Gal in triplicate samples in two hours, whereas, with the HPLC-based method, the same experiment would take approximately 22 hours.

A disadvantage is that the PCA is a discontinuous assay, the reactions are stopped by adding malachite reagents. However, when the PCA is compared to another potential continuous assay, where UDP is coupled to NADH oxidation via pyruvate kinase and lactate dehydrogenase,[84] less valuable materials are used. In addition, if the PCA is used for the

inhibitor characterisation, then the amount of false positive results is reduced due to fewer enzymes used and a longer detection wavelength (430 nm vs. 620 nm). The long detection wavelength is advantageous in inhibitor screening and characterisation as well as tolerance for DMSO and TX-100. Phosphate impurities and unspecific phosphatase inhibition can cause false positives in compound characterisation. However, the false positives can be avoided when the first screen is performed with another detection method (*i.e.* HPLC or LDA) or the assay is performed without coupling with to GT reaction. If the assay signal increases with increasing compound concentration then the compound likely contains inorganic phosphate and if the assay signal degreases with increasing compound concentration then the compound likely inhibits phosphatase or it is a real inhibitor. The feasibility of the phosphatase coupled assay for the HTS will be implemented in parallel with the ligand displacement assay and the PCA will be utilised for the inhibitor characterisation on the next chapter.

3.6 Experimental

3.6.1 HPLC

The HPLC experiments were carried out using a Perkin Elmer series 200 HPLC system with a series 200 EP with DAD. The analytes were separated with Alltech Alltima C18 column (3 µm particle size, 15 cm x 4.6 mm) in combination with a guard column (3 µ particle size, 7.5 × 4.6 mm, Alltech) and detected at 265 nm. The injection volume was 50 µL, the flow rate was 1.0 mL/min and the column oven was set to 30 °C. The data was acquired using Total Chrom Workstation version 6.2.0. All reagents used in this chapter are listed in App. 2.

The mobile phase A (buffer A) was consisted of 100 mM potassium phosphate and 8 mM TBAHS. The pH was set to 6.5 by mixing 0.064 mol of acidic and 0.036 mol of basic potassium phosphate (KH_2PO_4 and K_2HPO_4, respectively). The buffer B was a mixture of 70 % of buffer A and 30 % of MeOH (v/v). The mobile phases were filtered with 0.2 µm nylon filter and degassed in the vacuum prior the use.

The mixture of standard solutions of UMP, UDP and UDP-Gal at 25 µM each were prepared in water and used in the system suitability tests. The system suitability tests were performed with the gradient method (Table 3.11) and the Total Chrom Workstation performed the required calculations. The linearity of the method was performed by measuring the peak area of UDP and UDP-Gal at concentrations 1.0, 3.5, 10, 20, 25, 30, 50, 100 and 200 µM.

Table 3.11. The gradient of the optimised HPLC method used in the system suitability and K_m experiments

Time (min)	Buffer A (%)	Buffer B (%)
5	100	0
15	100-50	0-50
2	50	50
1	50-100	50-0
5	100	0

The LgtC catalysed reactions were carried out in 1.5 mL eppendorf tubes, in 50 mM MOPS buffer (pH 7.1), including of $MnCl_2$ (20 mM), BSA (1.0 mg/mL) and lactose (2.0 mM). UDP-Gal concentration varied depending on the experiment performed. LgtC was activated in 5.0 mM

DTT in MOPS buffer for 30 minutes at 37 °C, prior to use. The reactions were initiated by adding 0.63 nM of LgtC and incubated for 10 minutes at 37 °C. Total volume of the reaction was 100 µL. The reactions were stopped by placing the samples in dry ice and thawed prior to injection to the HPLC. The concentration of LgtC was chosen where the donor conversion to the product was 10 % or less by performing a calibration curve in various concentrations of LgtC.

All the following experiments were performed as described above.

- The influence of LgtC activation experiment was performed with activated and non-activated LgtC at UDP-Gal concentration of 200 µM.
- The reproducibility of the LgtC reaction was examined at fixed concentrations of all assay components at UDP-Gal concentration of 60 µM.
- The K_m experiment was performed at UDP-Gal concentrations of 1.6, 3.1, 6.3, 12.5, 25, 50, 100 µM. Controls were included at each UDP-Gal concentration in the absence of acceptor. UDP calibration curve at 1.0, 3.5 and 10 µM was included to each experiment. The K_m values were calculated by plotting the initial velocity over the range of UDP-Gal concentrations to GraFit 7.0.3.

3.6.2 Glycosidase coupled assay

A BMG Labtech PolarStar microplate reader with a fluorescence intensity detection mode was utilised for the development of the fluorescence based coupled assay. The experiments were carried out in black NUNC F96 MicroWell polystyrene plates. The fluorescence intensity was measured at 590 nm with excitation at 544 nm (resorufin λ_{ex} 573 and λ_{em} 590) and the flashes per well was set to 10.The gain was performed on the well with the highest concentration of resorufin (hydrolysed by β-galactosidase) on the linearity experiments and on the LgtC, and galactosidase coupled experiments, on the negative control well (no UDP-Gal, high fluorescence). The gain was set to 85 % on each experiment. All reagent used in this chapter are listed in App. 2.

The HPLC experiments were carried out with the HP Agilent 1050 series HPLC with DAD and the data analysis was performed with ChemStation software. Alltech Alltima C18 column (3 µm particle size, 15 cm x 4.6 mm) in combination with a guard column (3 µ particle size, 7.5 × 4.6 mm, Alltech) was used for the separation of the components of the LgtC catalysed reaction when Rgal was utilised as an acceptor.

Resorufin and Rgal stock solutions (50 mM) were prepared in DMSO. In this experimental chapter all dilutions and reactions were carried out in 50 mM HEPES pH 7.3 buffer, including 1.0 mM $MnCl_2$ and 1.0 mM $MgCl_2$, the total volume of the reactions was 200 µL and incubations were performed at 37 °C unless otherwise stated. In addition, the assay buffer was added to the controls instead of LgtC or UDP-Gal to keep the final volume at 200 µL.

3.6.2.1 Optimising the β-galactosidase concentration, calibration curve and stability of resorufin

The optimal β-galactosidase concentration was investigated at 20 µM of Rgal and β-galactosidase concentration from 0.01 U/mL to 8.2 U/mL (6 data points). The linearity of fluorescence of resorufin was investigated by hydrolysing Rgal. Concentrations of Rgal 50 µM were diluted by half, down to 0.8 µM (7 data points) and hydrolysed by 0.5 U/mL of β-galactosidase. The fluorescence intensity reading was taken after 10 minutes incubation and was utilised plotting the calibration curve. The stability of resorufin was investigated on the range from 0.8 to 50 µM (8 data points) by using standard resorufin solutions. Stability of resorufin was monitored for 60 minutes.

3.6.2.2 Coupling LgtC and β-galactosidase reactions

The investigations of whether Rgal was performing as an acceptor for LgtC, were carried out in the same HPLC setting and LgtC reaction conditions as those used as in chapter 3.6.1. Except Rgal was used as an acceptor at 100 µM and UDP-Gal concentration was 60 µM. The reactions were initiated by adding 0.05 or 0.005 U/mL of activated LgtC.

The investigations of the negative and positive controls on the fluorescence-based glycosidase coupled assay were performed at fixed concentrations of LgtC, β-galactosidase, UDP-Gal and Rgal at 0.1 U/mL, 0.5 U/mL, 50 μM and 6.3 μM, respectively. The influence of DTT was investigated by comparing negative controls that were performed with activated and non-activated LgtC. The negative control was carried out without LgtC or UDP-Gal and the positive control in the absence of β-galactosidase. LgtC initiated the reaction and the reaction mixture was incubated for 30 minutes and stopped by adding β-galactosidase. The gain was performed on the control well that contained Rgal and β-galactosidase (high fluorescence signal). Fluorescence intensity reading was monitored for 10 minutes after adding β-galactosidase.

The investigation of the progress of LgtC catalysed reaction on the glycosidase coupled assay was carried out at fixed concentrations of LgtC, Rgal and β-galactosidase at 0.1 U/mL, 6.3 μM and 0.5 U/mL, respectively. The UDP-Gal concentration was 5.0, 100 or 500 μM. The negative control was carried out without UDP-Gal and the positive control in the absence of β-galactosidase. The reactions were initiated by adding LgtC and stopped after 30, 45 or 60 minutes incubation by adding β-galactosidase. The fluorescence intensity reading was taken after 10 minutes incubation.

HPLC method for the analysis of the reaction progress of LgtC catalysed reaction when Rgal was used as an acceptor was modified from the previously developed method (Chapter 3.6.1) by increasing the content of MeOH (Table 3.12). The reaction progress was monitored in various incubation times by this modified HPLC method. The LgtC catalysed reactions were carried out in 50 mM HEPES buffer, including 1.0 mM of $MnCl_2$ at pH 7.1 and the concentration of UDP-Gal and Rgal was 200 μM. The reactions were initiated by adding 0.02 U/mL of activated LgtC (buffer for control) and stopped by placing the samples in dry ice after 30, 45 and 60 minutes incubations at 37 °C.

Table 3.12. Optimised HPLC gradient for the separation of UDP-Gal, UDP, Rgal and unknown product form LgtC catalysed reaction

Time (min)	Buffer A (%)	Buffer B (%)	MeOH (%)
2	98	2	0
5	98-60	2-40	0
1	60-75	40-0	0-25
12	75	0	25
1	75-98	0-2	25-0
5	98	2	0

3.6.2.3 LCMS method development

The method development for the separation of UDP, UDP-Gal, Rgal and the product of the LgtC catalysed reaction was performed with HP Agilent 1050 series HPLC with DAD and the data analysis was performed with ChemStation software. UDP and UDP-Gal were detected at 265 nm, and Rgal and the unknown product at 265 nm and 470 nm. The analytes were diluted in water:ACN (20:80) unless otherwise stated. Based on the requirement of the ESI in order to achieve optimal ionisation, the flow rate was set to 0.2 mL/min and ammonium acetate (NH_4Ac) was used as an additive. Additionally, the concentration of the additive was kept the same during the chromatographic run and the polarity of the mobile phase was varied.

Mobile phases were prepared by making NH_4Ac stock solution (100 mM) in ultrapure water at desired pH and diluted to the desired concentration with ultra-pure water and/or ACN. The mobile phases were filtered with 0.2 µm nylon filter and degassed in the vacuum prior the use.

Waters Atlantis HILIC silica column (3 µm particle size, 50 x 2.1 mm) was used for the analysis of mixture of UDP-Gal, UDP and Rgal (100 µg/µL each). 20 µL of the mixture was injected and the separation was started with isocratic run with 100 % of ACN:5 mM NH_4Ac (95:5) at pH 5.0. The column temperature was set to 30 °C.

Polymeric-based SeQuant ZIC-*p*HILIC column (5µm particle size, 150 x 2.1mm) was used for the analysis of mixture of UDP-Gal, UDP and Rgal (50 µg/mL each). 5 µL of the mixture was injected and the column temperature was set to 30 °C. Two mobile phases were used: buffer A consisted of ACN: 10mM NH_4Ac (pH 4) (95:5) and buffer B consisted of ACN: 10mM NH_4Ac (pH 4) (50:50) (v/v) and the separation was started with increasing content of buffer B. Various gradients at higher column temperature (at 60 °C) were examined to improve the separation of UDP and UDP-Gal (Table 3.13).

Table 3.13. Gradients examined with the ZIC-*p*HILIC column at 60 °C

Gradient	Time (min)	Buffer A (%)	Buffer B (%)
1	10	100	0
	1	100-27	0-75
	14	25	75
	1	25-100	75-0
	10	100	0
2	10	100-50	0-50
	1	50-30	50-70
	14	30	70
	1	30-100	70-0
	10	100	0
3	5	50	50
	3	50-35	50-65
	18	35	65
	1	35-50	65-50
	10	50	50
4 and 5[a]	5	100	0
	5	100-25	0-75
	15	25	75
	1	25-100	75-0
	10	100	0

[a]30 °C

Waters ACQUITY BEH Amide column (particle size 1.7µm, 150 x 2.1 mm) was used for the separation of mixture of UDP, UDP-Gal and Rgal (50 µg/mL). Injection volume was 5 µL. The separation was investigated with various gradients with increasing water content (Table 3.14). Two mobile phases were used: buffer A consisted of ACN:10mM NH_4Ac (pH 9) (95:5) and buffer B consisted of ACN:10mM NH_4Ac (pH 9) (50:50) or (30:70) (v/v). The gradient **a3** was used for the investigation of the peak shape in 10 %, 20 % and 30 % water in ACN, and on the reproducibility studies where the injection of mixture of UMP, UDP, UDP-Gal and Rgal (50 µg/mL each) was repeated six times.

Table 3.14. Gradients examined with the amide column at 30 °C

Gradient	Time (min)	Buffer A (%)	Buffer B (%)
a1[a]	15	100-0	0-100
	2	0	100
	1	0-100	100-0
	10	100	0
a2[b]	15	100-0	0-100
	2	0	100
	1	0-100	100-0
	10	100	0
a3[b]	20	100-0	0-100
	2	0	100
	1	0-100	100-0
	10	100	0

[a]Buffer B as ACN: 10mM NH_4Ac (50:50), [b] Buffer B as ACN: 10mM NH_4Ac (30:70)

Adaptation of the LgtC catalysed reaction to the amide HPLC method was carried out with the gradient **a3** (Table 3.14). The reactions were carried out in 1.5 mL eppendorf tubes in 50 mM HEPES buffer, including 1.0 mM of $MnCl_2$ at pH 7.1 and the concentration of UDP-Gal and Rgal was 200 µM. The reactions were initiated by adding 0.02 U/mL of activated LgtC and after 30 minutes incubation a 37 °C reactions were stopped by adding 800 µL ACN (total final volume 1.0 mL).

LC-ESI-MS/MS analysis was performed on a Thermo Scientific Accela LC system coupled to a Thermo TSQ Quantum Access MS/MS with an ESI source. Data acquisition was achieved using Thermo Finnigan Xcalibur software, version 4.1. Fraction patterns of the standard solutions of UMP, UDP, UDP-Gal and Rgal (10 µg/mL in mobile phase B, ACN: 10mM NH_4Ac (30:70)) were performed prior to the analysis of the product of the LgtC catalysed reaction. The reaction conditions and chromatographic method was set up as described above.

3.6.3 Phosphatase coupled assay

A multi-mode BMG Labtech PolarStar microplate reader with absorbance detection mode was utilised for the assay optimisation. The phosphatase coupled assay experiments were carried out in clear NUNC F96 MicroWell polystyrene microplates and absorbance was recorded at 620 nm. The K_m values were calculated by the initial velocity over the range of UDP-Gal concentrations to GraFit 7.0.3. All reagents used in this chapter are listed in App. 2.

All glassware used in the reagent preparation was washed with 1 M H_2SO_4 to ensure removal of phosphate and rinsed with ultrapure water. The malachite reagent A was prepared by diluting 6.3 N of sulphuric acid into ultrapure water, after cooling, 1.75 % w/v of ammonium molybdate tetrahydrate was added and then filled up to the required volume. The malachite reagent B was prepared by adding 0.35 % (w/v) of polyvinyl alcohol (PVA) to ultrapure water and dissolved by microwave heating. After cooling, 0.035 % of malachite green oxalate was added and filled up to the desired volume.

3.6.3.1 Optimisation with LgtC

All the assay components of LgtC reactions were diluted in 13 mM HEPES (pH 7.0) buffer and the reactions were carried out in mixture containing $MnCl_2$ (5.0 mM), CEL (1 mg/ml), CIP (10 U/ml), lactose (2.0 mM). DMSO (10 %, v/v) and TX-100 (0.01 %, v/v), if required. The reactions were initiated by adding UDP-Gal or UDP to the calibration curve/phosphatase control (Table 3.15).The reactions were incubated for 20 minutes at 30 °C and stopped by adding malachite reagent A and mixed thoroughly, and then malachite reagent B was added and again mixed thoroughly. The colour development was monitored for 30 minutes and the final reading was taken when the colour development was levelled out, approximately after 20 minutes. UDP calibration curve, which also served as a phosphatase control was carried out at 0, 0.78, 1.56, 3.13, 6.25 and 12.5 µM in all experiments. LgtC was activated with DTT prior to the experiments and the concentration of LgtC was selected where the UDP-Gal conversion to UDP was 10 % or less.

Table 3.15. Phosphatase coupled biochemical assay set up with LgtC

Assay step	Component	[Stock]	[reaction]	Volume of stock, µL
1	Lactose	10 mM	2 mM	30
	CIP	100 U/ mL	10 U/ mL	15
	$MnCl_2$	50 mM	5.0 mM	15
	CEL	10 mg/ mL	1.0 mg/ mL	15
	Buffer	5x	1x	15
2	LgtC/ buffer[a]	5x	1x	30
3	UDP-Gal/ UDP[b]	5x	1x	30
4	**20 min incubation at 30 °C**			
5	Malachite reagent A	5x	1x	30
6	Malachite reagent B	5x	1x	30
7	**colour development monitored at 620 nm for 30 minutes**			

[a]buffer added to the background hydrolysis and UDP calibration/phosphatase control wells, [b]UDP-Gal was added to all wells to initiate reaction and UDP was added to the UDP calibration/phosphatase control wells

The first experiment with LgtC was performed as presented in Table 3.15, in the absence and in the presence of CEL. The turnover of UDP-Gal was carried out at 20 µM and various LgtC concentrations (0, 0.32, 32 and 320 µM).

Enzymatic hydrolysis of UDP-Gal was performed in the absence of lactose at 0, 20 and 50 nM of activated LgtC and at 0, 3.13, 6.25, 12.5, 25, 50, 100 and 200 µM of UDP-Gal.

The influence of $MnCl_2$ concentration was investigated at 0, 1.0, 2.5, 5.0, 10, 16, 20 and 25 mM of $MnCl_2$. LgtC and UDP-Gal concentrations were fixed at 20 nM and 13 µM, respectively.

The influence of TX-100 for the colour development was investigated by performing UDP calibration curve at 0, 0.001 and 0.01 % (v/v) of TX-100. UDP calibration curve was carried out at 0, 0.78, 1.56, 3.13, 6.25 and 12.5 µM.

K_m was determined in various assay conditions: standard conditions (Table 3.15), in the presence of DMSO, in the presence of DMSO without $MnCl_2$ and in the presence of DMSO and TX-100. The experiments were carried out at 0, 3.13, 6.25, 12.5, 25, 50, 100 and 200 µM of UDP-Gal. The layout of the microplate of the ordinary K_m experiment is presented in Figure 3.24. The K_m was

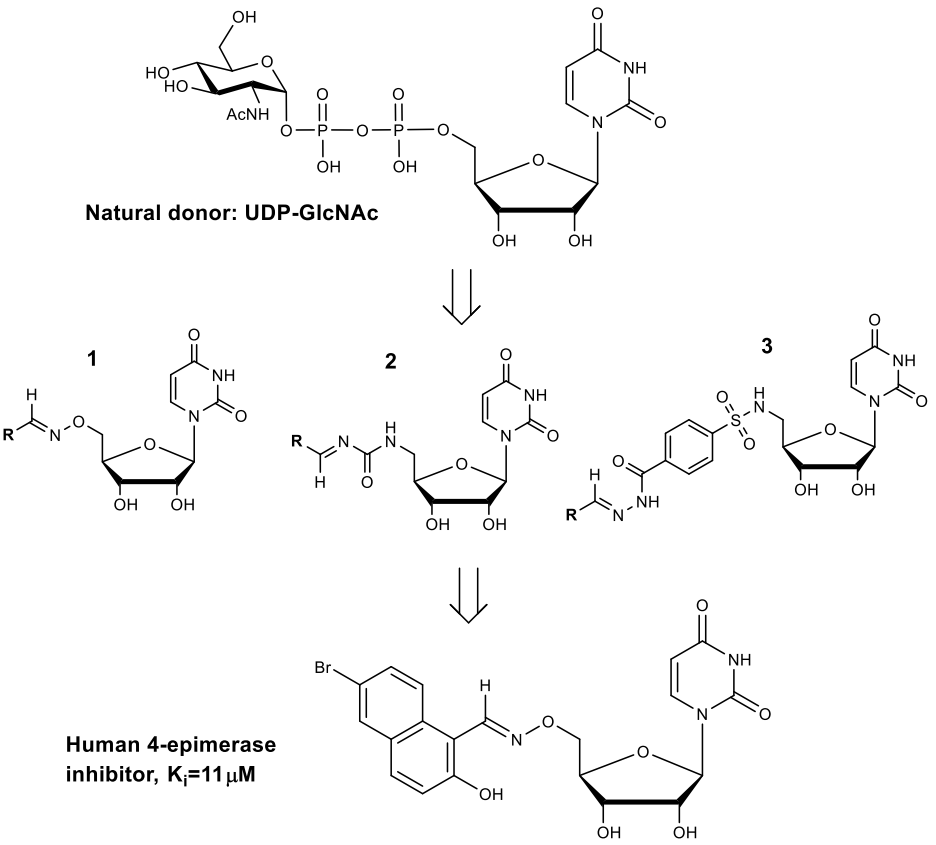
	1	2	3	4	5	6	7	8-12
A								
B		**M**	**M**	M+E	M+G1+E	M+G2+E	M+G3+E	→
C		**M+U1**	**M+U1**	M+E	M+G1+E	M+G2+E	M+G3+E	→
D		**M+U2**	**M+U2**	M+E	M+G1+E	M+G2+E	M+G3+E	→
E		**M+U3**	**M+U3**	M	M+G1	M+G2	M+G3	→
F		**M+U4**	**M+U4**	M	M+G1	M+G2	M+G3	→
G		**M+U5**	**M+U5**	M	M+G1	M+G2	M+G3	→
H								

Figure 3.24. Layout of the micro plate in K_m experiments. Columns 2-3: UDP calibration (phosphatase control), rows B-D from column 4: various UDP-Gal in the presence of LgtC, rows E-G from column 4: background at each UDP-Gal concentration in the absence of LgtC. Master mixture (M) contain lactose, $MnCl_2$, CEL and CIP, UDP at concentrations of 0, 0.78, 1.56, 3.13, 6.25 and 12.5 µM (U1-U5), UDP-Gal (G) and LgtC (E)

4 Compound library screening and compound characterisation

4.1 Introduction

High throughput screening (HTS) is a popular method in drug discovery that allows the screening of thousands of compounds against selected targets simultaneously. Compound libraries are collections of structurally diverse small compounds with undefined properties and biological activities[212] or focused subset of compounds that have proven biological activity to certain molecular targets.[213] Structurally diverse compound libraries have a broad range of biological activities and therefore, the possibility of discovering novel inhibitors is increased.[214] Focused compound libraries are designed often based on a certain protein family and the structures of the compounds are generally built on combinations of one scaffold and two or three side chains.[213] For example, Winans and co-workers synthesised over 1000 compounds based on the substrate of human UDP-GlcNAc 4-epimerase and from the subset, they identified a novel inhibitor with a K_i value of 11 μM (Scheme 4.1).[215]



Scheme 4.1. Schematic presentation of design of focused compound libraries. Uridine analogues (1-3) with alternative linkers mimic the structure of natural donor. Various functional groups (R) replace the natural sugar. Human 4-epimerase inhibitor discovered from the uridine analogue library.[215]

Drug-likeness of the compounds is important in the early stage of the drug discovery process and a number of compound collections have been produced by focusing the small molecule synthesis to drug-like compounds.[216, 217] The Lipinski rule of five is a general guideline to determine the drug-likeness and it evaluates solubility and cell permeability of the compounds to be an oral drug. Based on the Lipinski's rule, the drug-likeness of the molecules increases if the molecular weight is less than 500 Daltons, LogP (describes lipophilicity) is less than 5 and the molecule contains hydrogen bond donors and acceptors less than 5 and 10 respectively.[218] Recently, Bickerton and co-workers refined the Lipinski's rule of five and they called it quantitative estimate of drug-likeness (QED). Addition to the Lipinski's rule, QED included polar surface area of molecule, number of aromatic rings, rotatable bonds and groups causing toxicity.[219] The lipophilicity of the drug candidate molecule is the most important factor as the influence on absorption, metabolism, distribution and toxicity.[220]

Compound dependent assay interference can cause false positive and negative readouts in the HTS of small molecule libraries. Spectral properties of the small molecules can cause false results due to their fluorescence, quenching and inner filter absorbance effects.[221] The compound libraries generally contain a greater number of heterogeneous compounds that are fluorescent in the blue-green range and therefore, the assays that detect the fluorescence intensity in this range, are more prone to false positives.[222, 223] Compound's interference with the target enzyme is difficult to identify if the interference is concentration dependent and reproducible.[224, 225] Various small compounds are able to form 50-1000 nm particles (aggregates) that non-specially inactivate the target enzymes by sequestering the surface.[149] The compound aggregation is dependent on the assay conditions, properties of the compound and the target enzyme.[223, 226] However, the aggregation is reversible in the presence of a non-ionic surfactant such as TX-100.[149] In addition, the false positives can be caused by presence of reactive compounds that form covalent bonds with the target enzyme,[227, 228] redox cycling compounds such as strong reducing reagents (*i.e.* DTT) can reduce cysteine residues of the target enzyme and effect their activity and impurities can either interfere with the assay signal or with the target enzyme.[221, 228-230]

Generally, the HTS process involves more than two screens to eliminate the false positives and characterise the active compounds.[152] For example, Helm and co-workers performed two screens to eliminate false positives and a third assay was utilised in the hit characterisation.

Initially they screened nearly 50000 compounds against MurG, a GT that catalyses the transfer of GlcNAc from UDP-GlcNAc to Lipid I.[89] The HTS assay based on the fluorescence polarisation change upon the displacement of the fluorescence substrate. The primary screen was performed in a single compound concentration and 0.6 % of the compounds reached positive hit criteria. From the hits, 44 compounds with different structure scaffolds were selected and further analysed in the secondary screen. The secondary screen was based on monitoring a radiolabeled product and therefore, the different detection method eliminated possible false positives. 11 of the 44 compounds met the positive hit criteria and one of the compounds was characterised by a fluorescence based NADH coupled assay.[89]

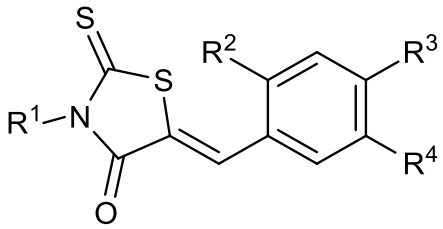
In this chapter, two compound libraries were screened against LgtC by the ligand displacement assay (LDA) which was developed in chapter 2 and the hits were characterised by the phosphate coupled biochemical assay (PCA) which was optimised in chapter 3. LgtC has been found to be involved in the biosynthesis of meningococcal LOS[43] which is a requirement for the survival of bacterial cells and increasing virulence of *H. influenzae*[113] makes LgtC a potential antibiotic target. A diverse compound library and a target-focused library with proven bioactivity against other carbohydrate active enzymes were screened in order to identify novel inhibitors and leads (compounds that have biological activity[152]). The leads could be further modified to improve their activity. The compound libraries are introduced in more details in chapters 4.2 and 4.3.

4.1.1 Compound library screening protocol

The aim was to use the LDA for a large compound library screening for the first time. Like many HTS assays, the LDA is based on a fluorescently labelled ligand (5FTUDP-Gal). The emission wavelengths of the fluorophore is short (430 nm) which increases interference with the screened compounds that emit with similar wavelength. The LDA was developed and optimised with non-fluorescent compounds such as UDP (chapter 2) and behaviour of screened compounds whether they quench or increase the fluorescence intensity of the fluorophore at desired wavelength was not investigated before. Therefore, a test experiment was performed before carrying out the compound library screening to evaluate the possible outcome from the actual screening.

Thiazolidinone compounds have been synthesised and characterised in house.[102] They have been reported to inhibit carbohydrate active enzymes such as UDP-galactopyranose mutase and dolichophosphate mannose synthase.[104, 231, 232] Pesnot and co-workers reported that one thiazolidinone bound to three GalTs (GTB, LgtC and α1,3-GalT) with similar affinity[104] which may indicate that the interaction between studied thiazolidinone and protein may not be specific. In addition, a systematic study of five-membered multiheterocycles, including thiazolidinones, suggested that the exocyclic sulphur atom offers interaction sites for polar interactions and hydrogen bonds which make them prone to bind to a large number of protein targets.[233] However, thiazolidinones were a good test case for the investigation of possible outcome of the coloured compounds from the library screening due to their aromatic systems and similar emission wavelengths as the fluorophore.[102]

Table 4.1. Structures of examined thiazolidinones



Compound	R^1	R^2	R^3	R^4
T1	-	-	-	-O-CH_3-Ph
T2	-Ph	-	-O-CH_3-Ph	-O-CH_3-Ph
T3	-CH_2-COOH	-CH_3	-	-
T4	-CH_2-COOH	-	-O-CH_3-Ph	-O-CH_3-Ph

Four thiazolidinones with various functional groups were selected for the test experiment (Table 4.1). The interference of fluorescence of the test compounds was examined in the presence of the fluorophore and compared to the pure fluorophore fluorescence signal. The results illustrated that with increasing concentration of the test compounds, **T1** and **T2** elevated and **T3** and **T4** quenched the fluorescence signal (Figure 4.1). At 100 µM of **T1** the fluorescence signal of the fluorophore was significantly increased and significantly quenched at 100 µM of **T3** and **T4**. The quenching can be caused by a high concentration of the compound and the light is unable to pass through the sample to cause absorption, or **T3** and **T4** may deactivate the excited fluorophore upon contact and heat is released instead of emitted photon (collisional quenching).[87] UDP was utilised as a positive control at 5.0 mM and the results also demonstrated that UDP did not have an influence on the fluorescence intensity signal.

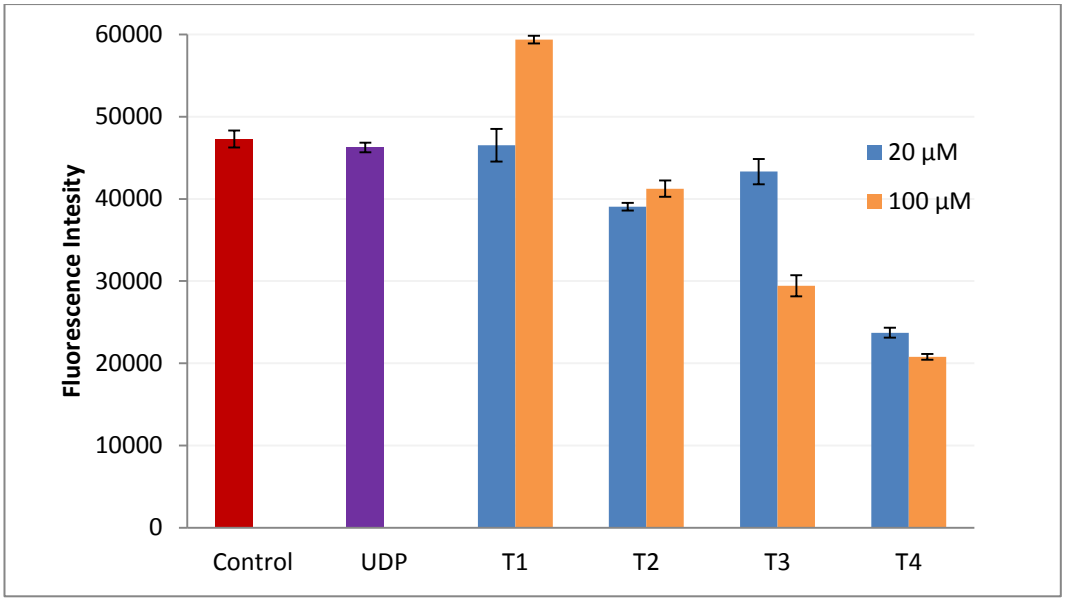


Figure 4.1. Tested thiazolidinones at 20 µM and 100 µM in the presence of the fluorophore. The control is pure fluorophore and UDP at 5.0 mM in the presence of the fluorophore. Fluorescence intensity was measured at λ_{ex}= 350 nm and λ_{em}= 430 nm which are relevant to the fluorophore, 5FTUDP-Gal

The quenching and the fluorescent properties of the tested compounds on the LDA had to be considered. Therefore, the ligand displacement was investigated with two compounds that increase and quench the fluorescence signal of the fluorophore. **T1** and **T3** were selected and tested against LgtC at 100 µM. The maximum fluorescence intensity was obtained at 5 mM of UDP (positive control) and the lowest fluorescence intensity in the absence of UDP (negative control). The fluorescence intensity change between the negative and the positive control (Figure 4.2, Δ1) is the assay window (a maximum fluorescence change) which corresponds to complete fluorophore displacement (i.e. 100 % of the fluorophore is displaced in the presence of UDP (5.0 mM)).

The fluorophore displacement of the tested compounds was unsuitable to determine in a similar way performed with UDP because the fluorescence intensity of the pure fluorophore was significantly changed in the presence of **T1** or **T3**. For example, because the fluorescence intensity was significantly quenched in the presence of **T3**, the sample signal (**T3** + fluorophore + LgtC) was lower than the negative control signal. Therefore, the same negative control was unsuitable for determining the percentage displacement. Consequently, the percentage change in the fluorescence intensity compared to the pure fluorophore (F) signal was used to calculate a negative control for both compounds by correcting the negative control signal (fluorophore + LgtC, F + E). For example, the fluorescence intensity increased

25 % in the presence of **T1** and quenched 38 % in the presence of **T3** at (Figure 4.2, Δ2 and Δ4, respectively). The negative control was then corrected for both test compounds depending on the percentage fluorescence increasing or quenching intensity assuming the change in fluorescence intensity is linear and remains the same during the screening experiment. The change in fluorescence intensity of the tested compounds was calculated between the corrected negative control and the sample signal (Figure 4.2, Δ3 and Δ5 for **T1** and **T3** respectively) which is related to the assay window achieved with UDP.

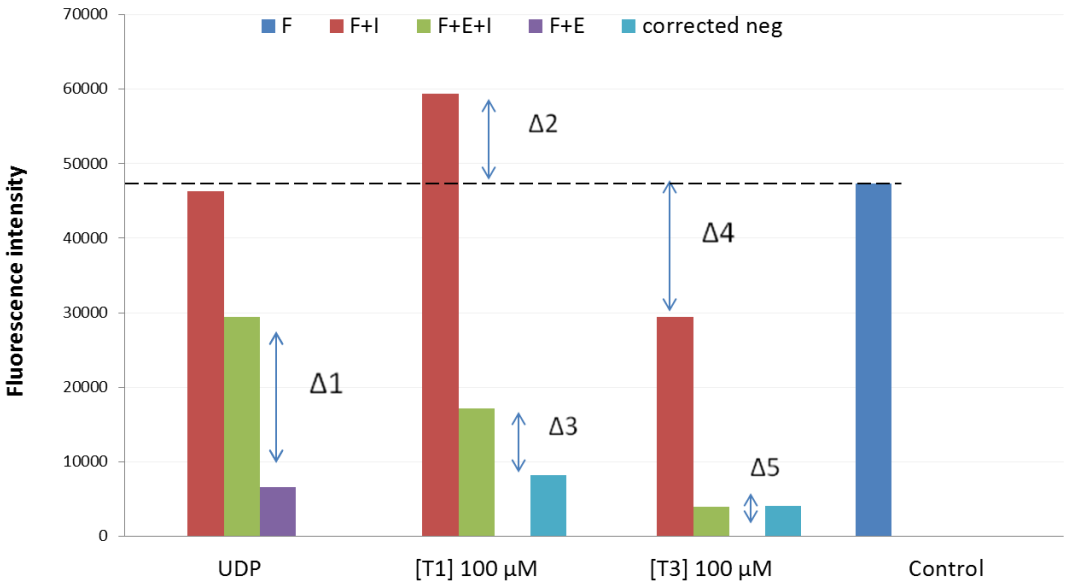


Figure 4.2. Fluorescence intensity (FI) comparison between the controls and **T1** and **T3** at 100 µM in the absence and in the presence of LgtC. **I** is binder: **T1**, **T3** or UDP, **E** is LgtC and **F** is the fluorophore (5FTUDP-Gal). Δ1 is the assay window: the maximum FI change between negative and positive control (UDP 5.0 mM), Δ2 and Δ4 is the difference in FI caused by the tested compounds compared to the control (F only), Δ3 and Δ5 is the difference in FI between the corrected negative control (corrected neg) and the FI signal of the test compounds in the presence of F

The percentage fluorophore displacement was calculated (Equation 6) by normalising the change in the fluorescence intensity between the measured sample (F_c) and the corrected negative signal (F_{corr}) to the maximum FI change between the negative control (F_{min}) and the positive control (F_{max}).

Equation 6

$$\%\text{Displacement} = \mathbf{100}\,\boldsymbol{x}\,\frac{(F_c - F_{corr})}{(F_{max} - F_{min})}$$

These results demonstrated how tested compounds may possibly behave and how to consider the fluorescence properties of the tested compounds in calculations. The negative control was corrected for each compound individually. The fluorescence intensity of the compounds was taken into account at this stage, therefore, the false negatives can be minimised. UDP behaves like a competitive inhibitor and it is non-fluorescent, therefore, it was utilised as a control in the compound screenings. UDP was introduced as a positive control at 5.0 mM and at concentration where 50 % of the fluorophore was displaced (IC_{50}) by UDP was set as a limit for positive hits.

The compound libraries, a diverse compound library and a target-focused library were first screened in a single compound concentration aiming to identify compounds that inhibit LgtC with the LDA. The second screen was performed in various compound concentrations (IC_{50}) also with the LDA to remove the potential false positives. Because the LDA measures only binding and does not provide information about the catalytic activity of the enzyme, the final screen was performed with a biochemical assay using a different detection mode. The phosphatase coupled assay, was then utilised to confirm and characterise the hit compounds and determine the kinetic parameters against LgtC (*i.e.* the inhibition constant, K_i).

4.2 Diverse compound library screening

A diversity-oriented synthesis (DOS) library was received from Dr David Spring. The compound library included several scaffolds and numerous building blocks of the final molecules. Some of these structures are illustrated in Figure 4.3, Figure 4.4 and Table 4.2. The diverse library ensures that a wide range of structural diversity is represented in the sample set. The library consisted of a structural diversity of compounds that did not have proven bioactivity activity against glycosyltransferases.

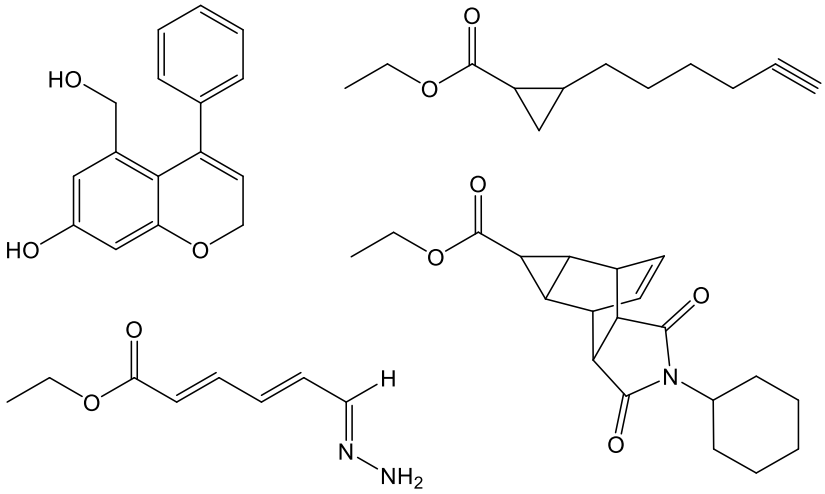


Figure 4.3. Examples of compounds included in the DOS library[217]

The compound screening demonstrated that the LDA is a rapid method as, in total 393 compounds were successfully screened in duplicate against LgtC in one day. Screening such a large number of compounds in one day is a great advantage. If the compound library was screened with the HPLC-based assay or the PCA then the screening would take approximately 40 or 5 days respectively. The Z'-factor of the average of 9 measured micro plates was 0.74. The compounds were tested at 100 µM and the hit rate was high, 8.1 % of the compounds displaced more than 50 % of the fluorophore. Typically the hit rate in the compound library screening is 0.01-0.1 %[224] and therefore, the high hit rate raised questions about false positives. Over 50 % of the hit compounds shared a bicyclic nitrogen heterocyclic scaffold (Figure 4.4 and Table 4.2). The hydroxyl group of the phenol ring of the compounds could possibly form chelates with manganese ion and cause high hit rate by inactivating LgtC. In addition, LgtC could be inhibited non-specifically by compound aggregation because the screening was performed in the absence of surfactants. The fluorescence of the compounds was unlikely to cause false negatives because the fluorescence of the compounds was taken in to account in the calculation (Equation 6).

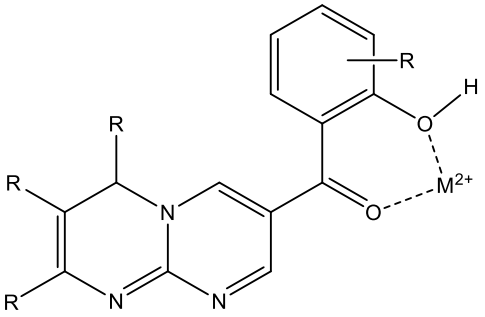


Figure 4.4. The general scaffold of the positive hits from the primary screen of the DOS library

4.2.1 Characterisation of the hits from the diverse compound library

The top 13 compounds that displaced the fluorophore more than 50 % were selected for subsequent investigations for the discovery of potential false positives. Firstly, the IC_{50} was determined with the LDA to ensure that the hits were real. The compounds were clearly insoluble in the assay buffer above 100 µM, therefore, the highest concentration of the compounds in the IC_{50} experiment was limited to 100 µM. The displacement of the fluorophore was dependent on the concentration of the compounds (Figure 4.5). However, accurate IC_{50} values were not achieved because a plateau was not reached in the concentration range used.

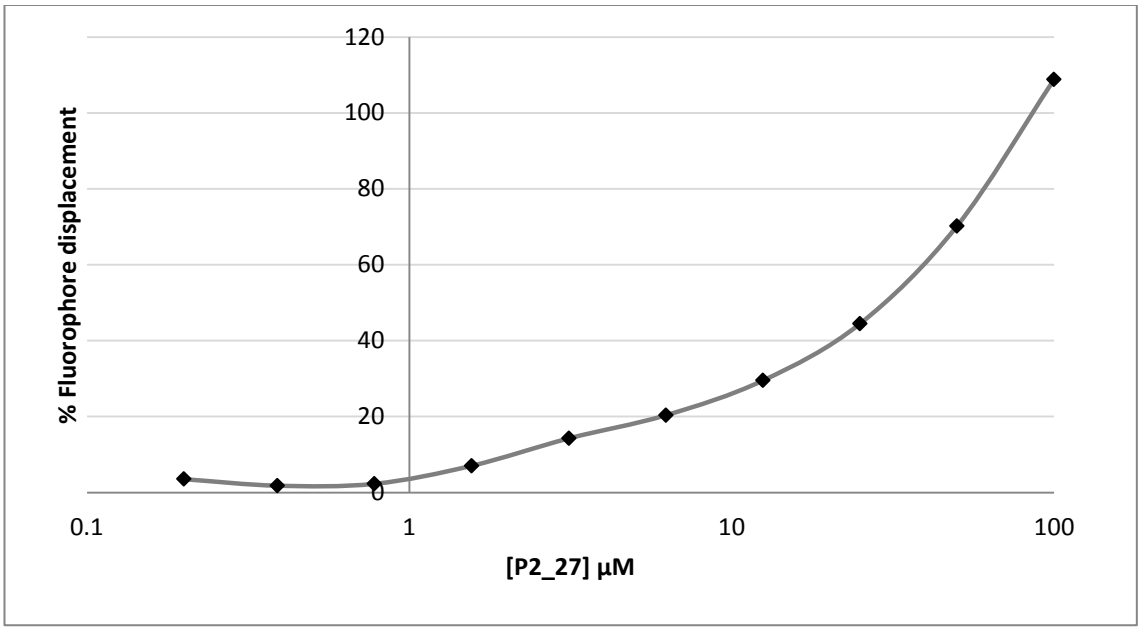
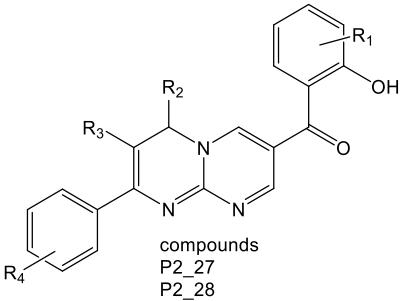


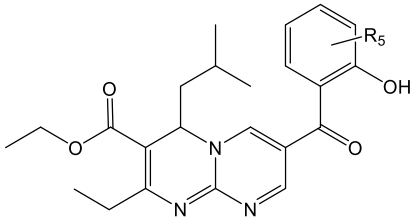
Figure 4.5. Fluorophore displacement with one of the hit compounds from the primary screen (P2_27) in various concentrations against LgtC. Experiment carried out in the presence of 10 % DMSO and fluorophore displacement monitored at λ_{ex}= 350 nm and λ_{em}= 430 nm

The HPLC-based method (Chapter 3.2.3) was a valid alternative method for investigating the activity of the compounds and to exclude the possibility of chelation because the turnover is monitored in the excess of manganese ion by the HPLC-based method. In addition, a different detection method could exclude false positives if the positive hits are caused by the fluorescence of the compounds. The percentage inhibition was determined at 50 µM by the HPLC method and the percentage inhibition was then compared with the fluorophore displacement results at the same concentration.

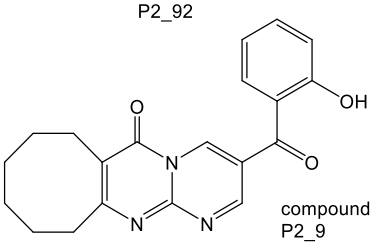
Table 4.2. Structures of the hit compounds and percentage inhibition at 50 µM against LgtC by HPLC and percentage fluorophore displacement by LDA



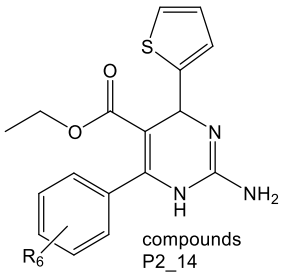
compounds
P2_27
P2_28
P2_31
P2_92



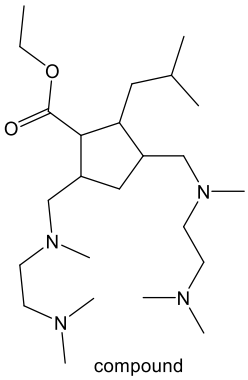
compounds
P2_83
P2_89
P2_93



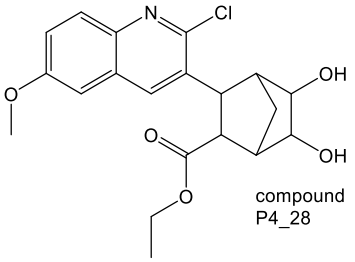
compound
P2_9



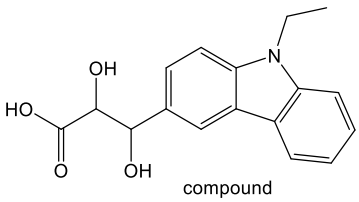
compounds
P2_14
P2_72



compound
P4_29



compound
P4_28



compound
P4_41

Com-pound	% INH HPLC[b]	% Displa-cement LDA[c]	R_1	R_2	R_3	R_4	R_5	R_6
P2_9	36	91	-	-	-	-	-	-
P2_14	33	64	-	-	-	-	-	*p*-Cl
P2_27	38	67	*m*-Cl, *m*-Cl	=O	-	-	-	-
P2_28	27	85	*m*-Cl, *m*-Cl	-T[a]	$-CO_2Et$	-	-	-
P2_31	37	81	-	$-CH_2CH(CH_3)_2$	$-CO_2Et$	*m*-OMe, *p*-OMe, *m*-OMe	-	-
P2_72	0	95	-	-	-	-	-	-
P2_83	29	69	-	-	-	-	-	-
P2_89	49	95	-	-	-	-	*m*-Cl, *p*-CH_3	-
P2_92	25	115	*m*-Br	=O	-	-	-	-
P2_93	32	98	-	-	-	-	*m*-Cl, *m*-Cl	-
P4_28	0	89	-	-	-	-	-	-
P4_29	20	49	-	-	-	-	-	-
P4_41	0	58	-	-	-	-	-	-

[a]Thiophene, [b]ralated to control (no compound added, 100 % activity), [c]related to 100 % fluorophore diplacement with excess of UDP

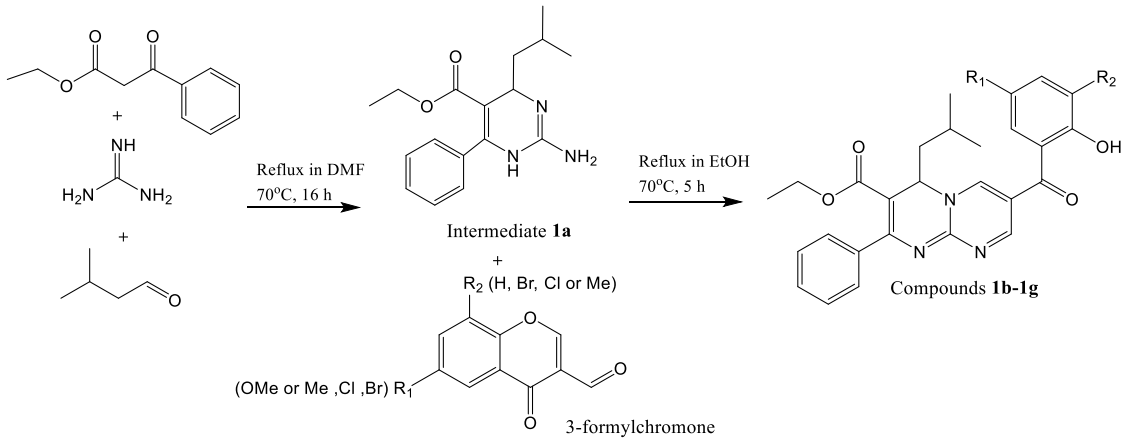
The observed percentage inhibition varied between 20 % and 49 % by the HPLC-based assay and the compounds **P2_72**, **P4_28** and **P4_41** did not show any activity against LgtC (Table 4.2). The chemical structures of **P2_72**, **P4_28**, **P4_41** and also **P2_29** are relatively different than the remaining hits. The percentage inhibition and percentage fluorophore displacement results cannot be directly compared because the HPLC-based assay monitor the turnover of the LgtC catalysed reaction and the LDA only the fluorophore displacement. Therefore, the observed activity of **P2_72**, **P4_28** and **P4_41** was possibly due to the ligand displacement assay conditions. The percentage fluorophore displacement varied between 49 % and 115%. Hypothetically, false positives can be observed if the compounds form chelates with Mn^{2+}. Mn^{2+} would then be unavailable and the fluorophore unable to bind to the donor binding site, resulting in a high fluorescence intensity signal on the LDA. This was an option because the LDA was performed without Mn^{2+} addition. However, the results obtained with the HPLC-based assay which was performed in the excess of Mn^{2+} (20 mM), indicated that the compounds inhibited the activity of LgtC. If the compounds inhibit LgtC by chelation then the activity of the compounds should not be observed in high Mn^{2+} concentration. Therefore, the activity of the compounds against LgtC was a real inhibition or the observed activity was due to non-specific binding, or due to reactivity of the compounds.

4.2.1.1 Synthesis of pyrimidopyrimidine compounds

Most of the hit compounds shared a bicyclic nitrogen heterocyclic scaffold (Table 4.2). This scaffold had the most potential as a target structure based on the high activity against LgtC in the LDA and in addition, the HPLC-based method indicated activity. More material was needed for enzymological studies to complete the hit compound characterisation and to explore the structure-activity relationship. Therefore, a small library was designed and synthesised based on the activity against LgtC and availability of commercial reagents.

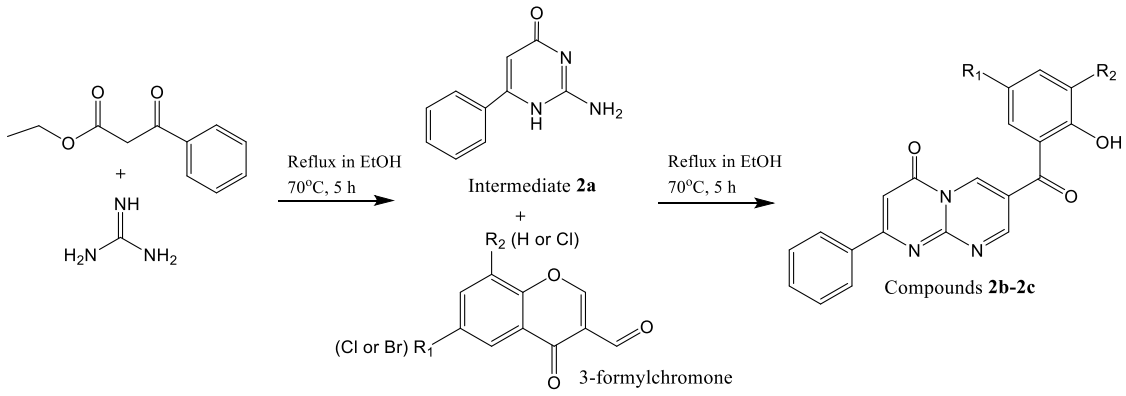
The synthesis route of these pyrimidopyrimidine (**PP**) compounds was simple and all the reagents were commercially available to synthesise the selected compounds. A small library of 10 compounds (Table 4.3) contained six compounds that were included in the primary screening (**1a**, **1b**, **1c (**same as **P2_27)**, **2a**, **2b** and **2c**) and the positive results also encouraged synthesise of four novel **PP** analogues (**1d**, **1e**, **1f** and **1g**).The benzoyl group attached to aminopyrimidine of the designed compounds remained unmodified and the

modifications were focused on the phenol ring on the other side of the heterocyclic scaffold by increasing the size with methane, chloride or bromide groups.



Scheme 4.2. Synthesis of **PP** analogues **1b** to **1g** via Biginelli-type reaction. The functional roups (R_1 and R_2) and yields are presented in Table 4.3

The synthetic route to the **PP** compounds was followed as previously published.[217, 234] The intermediate **1a** (Scheme 4.2) was prepared via a three-component Biginelli-type reaction[235] where β-keto ester (ethyl benzoylacetate), aldehyde (isovaleraldehyde) in the presence of guanidine generates a dihydropyrimidine in dimethylformamide (DMF). Whereas the reaction with β-keto ester and guanidine in EtOH generates amino pyrimidinone, the intermediate **1a** (Scheme 4.3). The final compounds were synthesised by modifying the intermediates **1a** or **2a** with an appropriate 3-formylchromone to form bicyclic nitrogen heterocyclic systems, **1b-1g** and **2b-2c**.



Scheme 4.3. Synthesis of **PP** analogues **2b** to **2c**. The functional roups (R_1 and R_2) and yields are presented in Table 4.3

The final compunds were purified by washing with hot EtOH or/and MeOH until pure compounds were observed in proton nuclear magnetic resonance spectroscopy (^{1}H-NMR). ^{1}H-NMR, ^{13}C-NMR and MS were utilised for the characterisation (Chapter 4.5.4) and the purity of the compounds was determined by HPLC ensuring that possible impurities did not interfere the assay performance: the ligand displacement assay or the phosphatase coupled assay on the enzymological studies. The absorption spectrum of the synthesised compounds was measured to determine λ_{max} prior to the purity determination (Chapter 4.5.5).

Table 4.3. The functional groups (R_1 and R_2, see the positions from the Scheme 4.2 and Scheme 4.3), yields and purities of the synthesised **PP** analogues

Compound	R_1	R_2	Yield (%)	Purity[c] (%)	λ_{max} (nm)
1a	-	-	49	-	-
1b	-Br	-	46	96	420
1c[a]	-Cl	-Cl	39	95	420
1d[b]	-Cl	-	55	99	380
1e[b]	-Me	-Me	51	96	380
1f[b]	-Br	-Br	43	41	380
1g[b, d]	-OMe	-	-	-	-
2a	-	-	36	100	310
2b	-Br	-	2	-	-
2c	-Cl	-Cl	26	24	310

[a]Same compound as P2_27, [b]novel compounds, [c]purity determined by HPLC, [d]product synthesis not achieved

Both intermediates (**1a** and **2a**) and six final products (**1b**, **1c**, **1d**, **1e**, **1f** and **2c**) were successfully synthesised up to 55 % yield (Table 4.3). Impurities of the final compounds were poorly soluble in used solvents (EtOH and MeOH) which hampered the purification step. The final compounds (**1b-1g** and **2b-2c**) were not soluble in hot EtOH or/and MeOH and washing was carried out in these solvents. Pure compounds were achieved by repeating this washing step, except with **1f** and **2c**. The reaction product (**1g**) was not formed during the reflux (precipitation was not formed), and a mixture of compounds was observed by thin layer chromatography (TLC). **2b** was successfully synthesised and pure compound was observed by NMR, however, most of the product was lost in the purification step and sufficient amount of the compound was not achieved for further analysis. The intermediate **1a** was not UV-active and the purity determination was impossible with HPLC, however, the ^{1}H-NMR analysis indicated a high purity product. The pure synthesised compounds: **1a**, **1b**, **1c**, **2a** and the novel compounds: **1d** and **1e** were chosen for the completion of hit compounds characterisation by the ligand displacement and the biochemical assays.

4.2.1.2 Determination of IC_{50} by ligand displacement assay

The preliminary results indicated that the inhibition by the hit compounds was concentration dependent (IC_{50} experiment) and the results in the chapter 4.2.1 indicated that the inhibition of LgtC was not due to the compounds ability to form chelates with the manganese ion. A potential explanation for concentration dependent inhibition was an aggregation-based inhibition. The aggregation-based inhibition is caused by molecules that can form small particles which inhibit proteins by sequestering LgtC at micromolar concentrations in aqueous solutions[236] and is dependent on a microenvironment, such as buffer concentration and pH, concentration of DMSO and surfactant.[237-240] Shoichet's group has investigated non-specific inhibition by molecular aggregators and they have reported that the aggregation-based inhibition can be reversed or prevented by adding a surfactant to assay buffer.[236] Therefore, the possibility that the hit compounds formed small particles which non-specifically inhibited LgtC was a potential explanation because the compound screening was performed in the absence of a surfactant.

Firstly, in order to investigate possible compound aggregation, the IC_{50} experiment was carried out in the presence and in the absence of TX-100 with the LDA. In theory, IC_{50} value of an aggregate-based inhibitor increases in the presence of surfactant and will have no effect for a real competitive inhibitor.[209] As mentioned earlier, the **PP** analogues were poorly soluble in the assay buffer above 100 μM, therefore, the highest concentration of the compounds in the experiment was 100 μM.

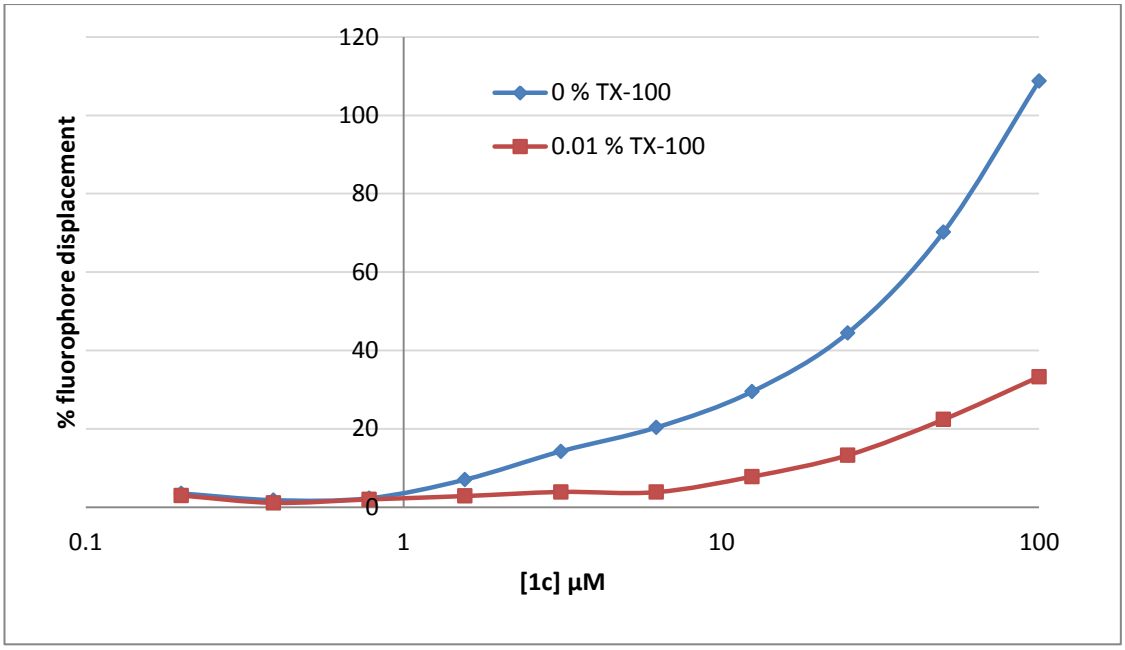


Figure 4.6. Determination of IC_{50} for **1c** in the presence and in the absence of TX-100 by the LDA

The plateau of the curve was not achieved in the concentration range used and the IC_{50} value was not determined, however, a significant effect of surfactant was observed (Figure 4.6). The fluorophore displacement at 3 µM and above of **1c** was four fold reduced in the presence of TX-100 indicating that the observed compound activity is possibly aggregation-based.[209] Additionally, the fluorophore displacement rose rapidly with the concentration in the absence of surfactant. The steep dose-response curves are reported to be an indication of aggregation-based inhibition. Shoichet has reported that the aggregation-based inhibitors can increase inhibition from 10 to 90 % less than 10-fold concentration range whereas a single-site inhibitor can do the same over an 81-fold concentration range.[241]

4.2.1.3 Determination of K_i by phosphatase coupled assay and investigation of aggregation-based inhibition

The determination of critical IC_{50} values which could have been used for comparing and characterising **PP** analogues was not yet achieved. Therefore, the phosphatase coupled biochemical assay (chapter 3.4.1) was introduced as an alternative method for the determination of K_i and aggregation-based inhibition of the compounds. The PCA is based on a different detection technique to the LDA. The advantage is that the PCA monitors the

turnover of the enzymatic reaction, providing information about the activity of LgtC in the presence of the tested compounds. Based on the same principle as the IC_{50} experiment, the K_i of an aggregate-based inhibitor increases in the presence of surfactant. Moreover, if the **PP** analogues are molecular aggregators, the inhibition is expected to be non-selective and they should be active against other enzymes.[242]

In order to investigate whether the inhibition was aggregation-based, the K_i was determined in various TX-100 concentrations against LgtC and β1,4-GalT. Additionally, the influence of Mn^{2+} was examined. The K_i experiments were carried out with all synthesised **PP** analogues (**1b-1g**), including the intermediates (**1a** and **2a**), the first scaffold of the **PP** structure which includes a phenyl attached to aminopyrimidine or aminopyrimidone (phenyl pyrimidine). In addition, the second scaffold which includes a benzoyl group attached to aminopyrimidine (benzoyl pyrimidine, **BP**, Figure 4.7) and one active compound from the compound screening (**P2_93**, from the batch received and not synthesised) were included for investigations of the structure-activity relationships and compound aggregation.

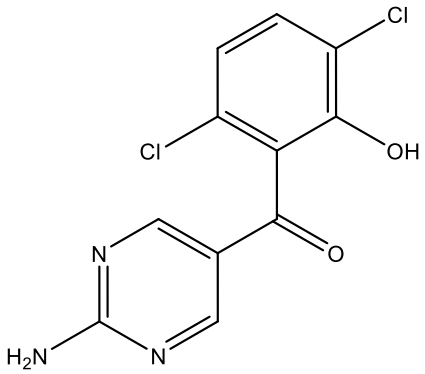


Figure 4.7. Structure of benzoyl pyrimidine, **BP**

Prior to the K_i determination the function of the PCA required investigation. A phosphatase catalysed reaction (Chapter 3.4.1) of the PCA is critical for the formation of inorganic phosphate that is proportional to the reaction product, UDP. Therefore, a control test was carried out to ensure that the phosphatase was not inhibited by the compounds of interest. The control experiment was investigated by titrating various concentrations of the **PP** analogues at fixed concentration of UDP and phosphatase. The results indicated that the same amount of inorganic phosphate/UDP was formed in the presence of the **PP** compounds. Because the phosphatase was unaffected by the **PP** analogues (Figure 4.8) and the PCA was suitable for K_i determination.

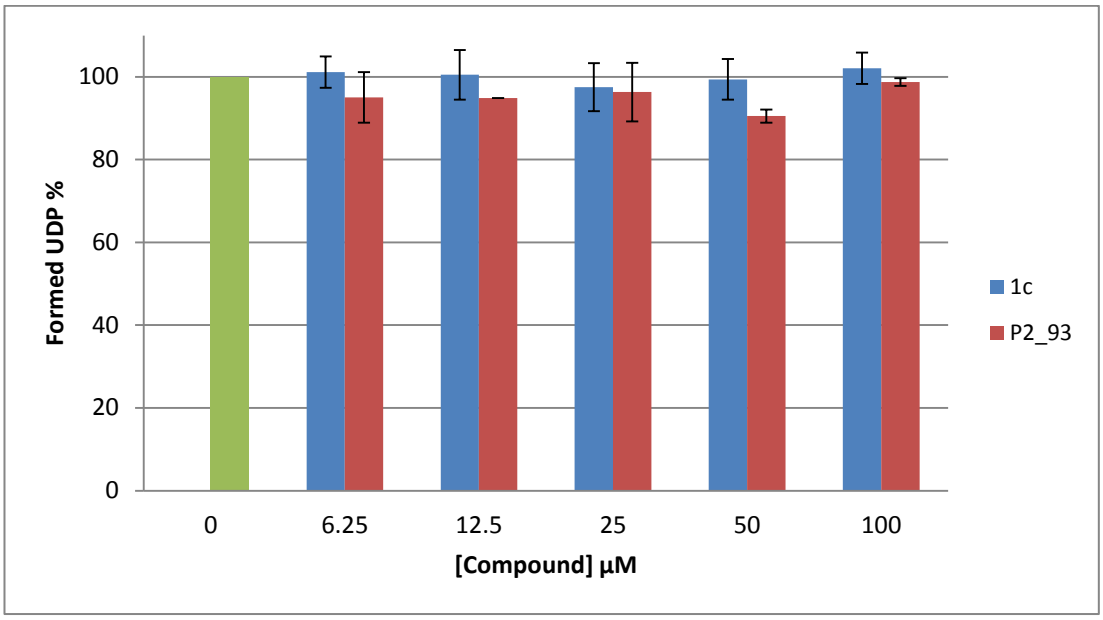


Figure 4.8. Phosphatase control by the PCA. The functionality of the phosphatase was not affected by UDP formation in various concentrations of **1c** or **P2_93**. Experiment was carried out at fixed concentrations of UDP (6.25 µM) and phosphatase (10 U/mL)

All the experiments were performed without addition of Mn^{2+} aiming to keep the same assay conditions such as in the LDA. The influence of Mn^{2+} was examined by determining K_i in the presence (5 mM) and in the absence Mn^{2+} for **P2_93** and **1c**. The K_i was 8.2 µM and 7.5 µM in the presence and 12 µM and 4.5 µM in the absence of Mn^{2+} for **P2_93** and **1c** respectively (Table 4.4). The unchanged K_i values indicated that the activity of LgtC was not influenced by the Mn^{2+} concentration. In addition, the results supported the hypothesis that the **PP** analogues were not chelating agents.

The K_i values of the **PP** analogues increased with increasing TX-100 concentration (Table 4.4). For example 10-fold increase in the K_i value of **1b** was observed when TX-100 increased from 0 % to 0.01 % with LgtC. In addition, the **PP** analogues inhibited β1,4-GalT with the similar concentration range than LgtC. The intermediates (**1a** and **2a**) and the second scaffold in the **PP** structure (**BP**) did not have activity against LgtC or β1,4-GalT.

Table 4.4. K_i values (μM) determined at 0, 0.001 and 0.01 % of TX-100 against LgtC and β1,4-GalT by PCA

Compound	% Fluorophore displacement[a]	LgtC			β1,4-GalT	
		K_i	K_i (0.001 % TX-100)	K_i (0.01 % TX-100)	K_i	K_i (0.01 TX-100)
P2_93	126	12 (8.2)[b]	-	-	6.6	-
1b	56	3.8	4.9	39	7.7	108
1c	108	4.5 (7.5)[b]	14	55	7.7	194
1d	-	1.9	13	57	19	122
1e	-	10	6.8	110	5.5	75
1a	-	116	>200	>200	>200	-
2a	-	>200	>200	>200	>200	-
BP	-	>200	-	-	>200	-

[a]Determined by the LDA at 100 μM, [b]Determined in the presentence of 5.0 mM $MnCl_2$

The results indicated that the **PP** analogues were surfactant sensitive and they were active against both, LgtC and β1,4-GalT suggesting that the inhibition is not selective. Both enzymes are possibly inhibited non-specifically by molecular aggregators that can physically sequester the surface of the protein and can cause protein unfolding.[236] A diverse range of compounds have been reported to aggregate and the prediction of the compound aggregation based on the structure is theoretically impossible.[223, 242] However, conjugated, lipophilic, poorly soluble in aqueous solutions and phenol group containing compounds are common features for reported aggregation-based inhibitors.[227] **BP**, which represents the second scaffold in the **PP** structure, has a phenol group which could be prone to aggregation, was used as a control to examine the aggregation and if a real inhibitor was hidden in the **PP** structure. However, no inhibition of LgtC or β1,4-GalT was observed by **BP** or the intermediates which are less aggregator-like compounds than the **PP** analogues. The results indicated that individually the two different scaffolds (**BP**, **1a** or **2a**) do not form aggregates, however, when these two scaffolds are combined, the **PP** analogues, aggregates are formed.

Overall, the results strongly suggested that **PP** analogues (**P2_93**, **1b**, **1c**, **1d** and **1e**) are aggregate-based inhibitors based on the surfactant sensitivity and activity against two different GTs. Additionally, all the **PP** analogues are highly conjugated and hydrophobic molecules, therefore, they could be molecular aggregators.

4.2.1.4 Determination of particle size by DLS

The results in the previous chapter indicated that the **PP** analogues were molecular aggregators due to their surfactant (TX-100) sensitivity and non-specific inhibition behaviour. To further investigate the aggregation behaviour of the **PP** analogues an alternative technique for detection of the particle size was required. Transmission electron microscopy and dynamic light scattering (DLS) methods have been used to determine a particle sizes in solution.[238, 243, 244] McCovern and co-workers investigated aggregation-based inhibition and they visualised that the aggregates formed by promiscuous compounds sequester enzyme by transmission electron microscopy.[149] DLS is a widely used technique for the investigation of particle size of aggregates and recently DLS has been coupled with plate readers allowing a particle size analysis parallel with high throughput screens.[243, 245]

DLS measures the intensity fluctuations of the scattered light arising from the Brownian motion of spherical particles in solution.[246] The Brownian motion is related to the particle size: larger particles move slowly and small particles quickly. The diffusion constant of the particles is measured and the size of the particles is calculated according to the Stokes-Einstein relation[246] ($D = kT/6\pi\eta r$, where k is Boltzmann's constant, T is absolute temperature, η is viscosity and r is radius of the spherical particle). The Brownian motion is dependent on temperature, viscosity and concentration of the solution.[246] In addition refractive index of the molecule of interest must be different than the diluents.

The particle size of the synthesised **PP** analogues (**1a**, **2a**, **1b**, **1c**, **1d** and **1e)** was determined by DLS. Sample diluent was copied from the LDA and the PCA, thus the compounds were dissolved in 13 mM HEPES pH 7.0 in final 10 % DMSO (v/v). The refractive indexes of **PP** analogues were between 1.58 and 1.67 (experimental chapter 4.5.6) and dissenting from the diluent (1.346[247]). Initially, the optimum concentration of the **PP** analogues for the particle size analysis was investigated. The particle size of **1b** was measure at 10 µM, 50 µM and 100 µM covering the concentration range used in the LDA and the PCA experiments. The particle size remained the same between 10 µM and 50 µM, and at 100 µM the particle size decreased significantly (Figure 4.9). DLS measures a single scattering events and possible multiple chattering events occurred at high sample concentration causing decrease in particle size. During the measurements of **PP** analogues, solubility problems were observed and these results indicated that at 100 µM the compounds were not completely dissolved.

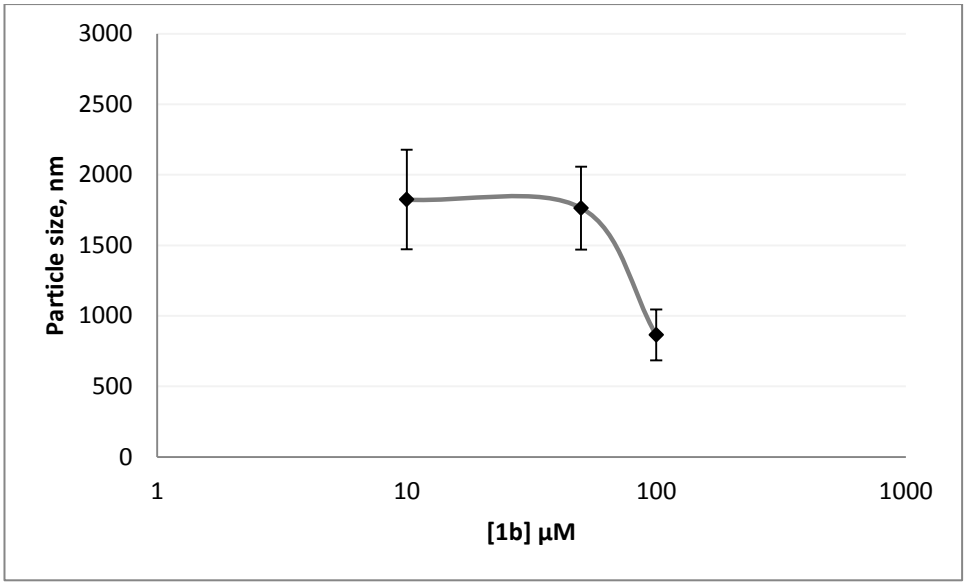


Figure 4.9. Particle size of **1b** at 10 µM, 50 µM and 100 µM. Optimum concentration for DLS experiments was found between 10 µ and 50 µM of **1b**

The relationship of the compound concentration and the aggregate formation has been investigated[240, 244] and the researchers demonstrated that each aggregating compound has a critical aggregating concentration and above the concentration more aggregates are formed, however, the particle size remains the same. Based on these investigations, the optimum concentration for the particle size experiments was chosen at 25 µM where particle size of **1b** was assumed as remaining the same.

Table 4.5. Average particle size of **PP** analogues (*n*=3) in 13 mM HEPES buffer (pH 7.0) including DMSO (10 % (v/v))

Compound	Size, nm	PDI[a]	Intensity kcps[b]
1a	217 ± 3	0.231 ± 0.027	214 ± 1
2a	84 ± 2	0.183 ± 0.022	87 ± 6
1b	1801 ± 123	0.712 ± 0.033	164 ± 2
1c	178 ± 6	0.354 ± 0.022	237 ± 2
1d	171 ± 4	0.339 ± 0.013	204 ± 2
1e	126 ± 1	0.224 ± 0.032	286 ± 6

[a]Polydispersity index, [b]Kilocounts per second

The results indicated that small particles were formed at 25 µM of each compound (Table 4.5). The average particle size was approximately 200 nm, which indicate that the particles may be aggregates. The reported particle size of aggregates is between 100 and 1000 nm.[238, 242] The particle size of **1b** was larger than the other tested compounds. The structure of **1b** and **1d** differs only from in functional group, **1b** has bromide and **1d** chloride on the same

position on phenol ring. The high polydispersity index indicated that the solution contains a variable size of particles, therefore, an alternative explanation is that a non-uniform solution of **1b** contains more large than small particles. Scattered light intensity remained high for all compounds, indicating a high number of molecules in solution. The intermediates (**1a** and **2a**) did not inhibit LgtC or β1,4-GalT (Table 4.4) by the PCA and they were not expected to form particles. However, the particle size of **1a** was in the same range as the final compounds (**1c-e**) and **2a**, which is smaller and more polar than **1a**, formed particles, but they were smaller than **1a** (Table 4.5). Possibly, the shape of the particles is unfavourable to sequester LgtC and β1,4-GalT compared to the final compounds (**1c-e**) that inhibited both enzymes.

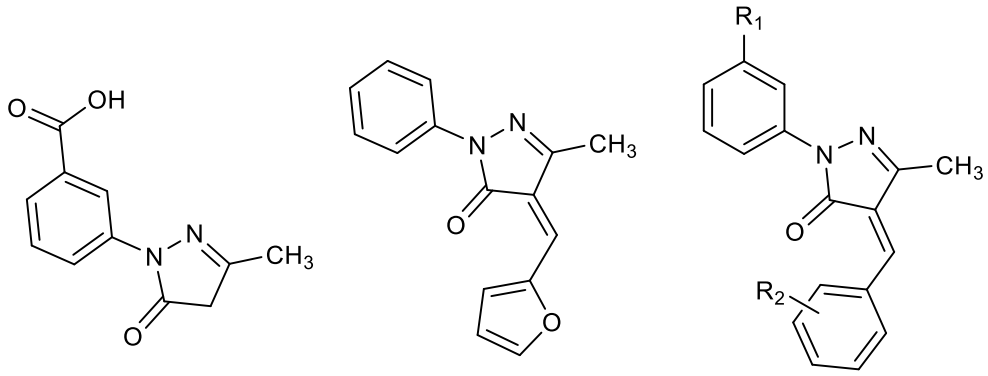
The results suggested that the compounds (**1b-e**) form aggregates which can non-selectively inhibit the activity of LgtC and β1,4-GalT. The results also indicated that the compounds were not completely soluble in the assay buffer at 25 µM. Therefore, the non-soluble particles may disturb the measurements. In addition, the precipitation of the compounds together with molecular aggregators may cause the false positive results in the compound screening and the biochemical assays.

4.3 Target-focused library screening

In recent years, target-focused libraries have been designed for improving the hit rates in HTS. The compound libraries have been designed based on the protein families[248] and the three-dimensional structure of the target protein.[249] Pyrazolones have been identified as novel GT inhibitors by virtual screening of 5 million commercially available compounds by targeting the acceptor binding site of heptosyltransferase (WaaC).[250] The enzymological studies demonstrated that pyrazolones were competitive inhibitors against WaaC[250] and the catalysis of the transfer of heptose to KDO (3-Deoxy-D-manno-oct-2-ulosonic acid) residue on the LPS of *E. coli*[251] was inhibited in the micromolar range. Pyrazolone derivatives have also been identified as kinase[252, 253] and HIV-1 integrase[254] inhibitors and potential antibacterial agents.[255, 256] A pyrazolone scaffold was chosen as a target structure for the discovery of inhibitors against LgtC, based on the pyrazolones activity against WaaC and the biological function of WaaC on the Gram-negative bacterial cell wall biosynthesis.

A small pyrazolone library was available in house and a set of 21 pyrazolone analogues (Table 4.6) was screened by the LDA and the PCA. The outcome of the compound screenings was used for the investigation of correlation between percentage inhibition and the percentage fluorophore displacement which was related to the maximum fluorophore displacement with UDP (see Chapter 4.1.1). Further characterisation of the positive hits identified from the screening were characterised by the PCA. TX-100 (0.01 %, v/v) was introduced to both assays to prevent possible false positives due to compound aggregation. The compounds were tested in duplicate at 25 µM and the screen was divided into two days. Z' factor average of two microplates was 0.91 with the LDA.

Table 4.6. Structure of the screened pyrazolones analogues



p1 **p2** **p3 - p21**

Compound	R_1	R_2
p3	-CO_2H	-
p4	-	*m*-OH, *p*-OH
p5	-CO_2H	*p*-OBn
p6	-CO_2H	*m*-Br, *p*-OCH_3
p7	-CO_2H	*p*-$N(CH_3)_2$
p8	-	*m*-OBn, *p*-OBn
p9	-	*o*-OH, *p*-OH
p10	-CO_2H	*m*-Br, *p*-F
p11	-	*m*-Br, *p*-F
p12	-	*m*-OCH_3, *m'*-OCH_3, *p*-OH
p13	-	*p*-Cl
p14	-	*p*-CN
p15	-	*m*-Br, *p*-OCH_3
p16	-	-
p17	-	*m*-CF_3
p18	-	*o*-CF_3
p19	-	*p*-CF_3
p20	-	*p*-Ph
p21	-	*p*-SO_2CH_3

The results from the compound screening showed that the percentage fluorophore displacement was significantly lower for several compounds by the LDA than the percentage inhibition by the PCA (Figure 4.10). The maximum percentage fluorophore displacement observed by the LDA was 27 % whereas 9 out of 21 (43 %) compounds inhibited LgtC more than 50 % by the PCA. The results at this stage indicated that the compounds are low affinity inhibitors by the LDA and false positives by the PCA or the percentage inhibition is real and the compounds are not competitive inhibitors. However, there were no suitable controls that could be used in both assays to prove which results were correct.

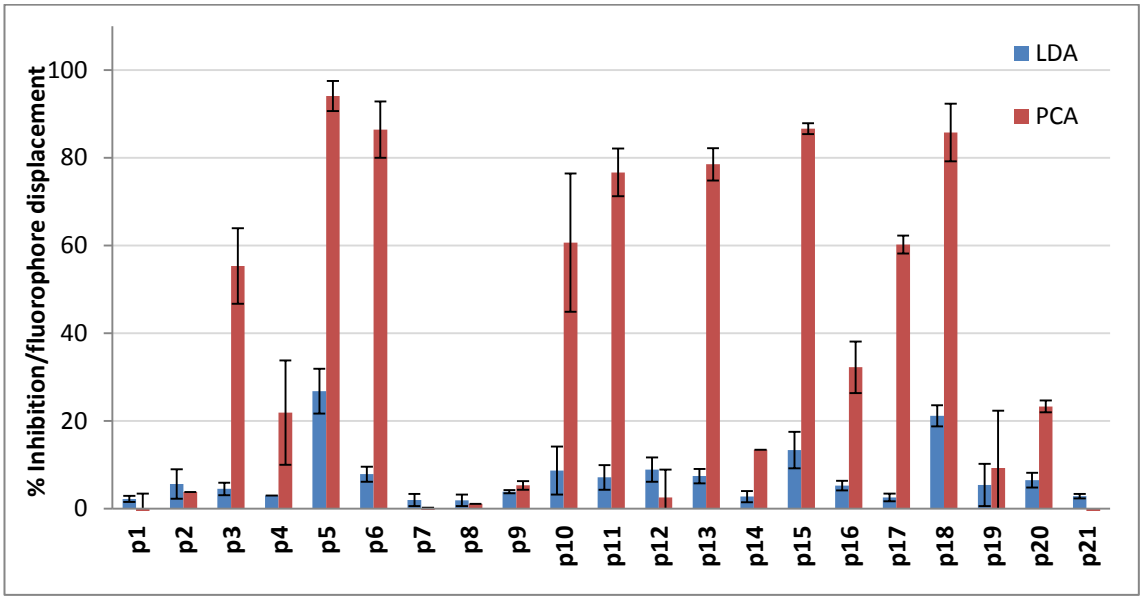


Figure 4.10. Percentage inhibition/fluorophore displacement at 25 µM of the compound against LgtC obtained by the LDA and the PCA

The percentage inhibition was used for the investigation of the structure-activity relationship by assuming that the results obtained with the PCA were real. The benzylidene moiety attached to the methylpyrazolone scaffold was essential for the inhibition. In the absence of benzylidine (**p1**) or when benzylidine moiety was replaced with a vinyl furan group (**p2**), a limited inhibition was observed. These preliminary results suggested that, a carboxylic acid group is not relevant to the inhibition when non-carboxylated and carboxylated compounds, **p6** and **p15** or **p10** and **p11** respectively, were compared. In contrast, the position of the functional group and the characteristics of the functional group of the aryl (R_2) ring seem to affect to inhibition. High polarity and large size of the functional groups seem to decrease inhibition.

The compound screening results obtained by the LDA and the PCA did not correlate as assumed. Therefore, further investigations were required in order explain the difference in the results. The type of inhibition was of particular interest as if the compounds inhibited LgtC with mixed mode then the low percentage fluorophore displacement and the high percentage inhibition could be explained.

4.3.1 Enzymological characterisation of hits from the focus-target library

Nine compounds were identified on the screening assay by the PCA that inhibited LgtC more than 50 %. However, compounds which displaced more than 50 % of the fluorophore (related to 100 % fluorophore displacement with UDP) were not identified by the LDA. Because the results obtained with these two assays did not correlate and the fluorescence intensity was partly recovered (7 to 27 %) by the LDA, it was suggested that the inhibition may occur in a mixed mode.

Reversible enzyme inhibition can be competitive, un-competitive or mixed. Competitive inhibitors binds specifically to the active site of a free enzyme and compete with the substrate.[257] Un-competitive inhibitors bind at a separate site than the substrate and they bind only to the enzyme-substrate complex.[257] Mixed inhibitors (and non-competitive) bind either to free enzyme or enzyme-substrate complex (Figure 4.11) with the same affinity (non-competitive inhibition) or different affinity (mixed inhibition).[257] Mixed inhibition is a mixture of partial competitive and pure non-competitive inhibition, and in the presence of inhibitor, some of the enzyme is always in non-active enzyme-substrate-inhibitor complex.

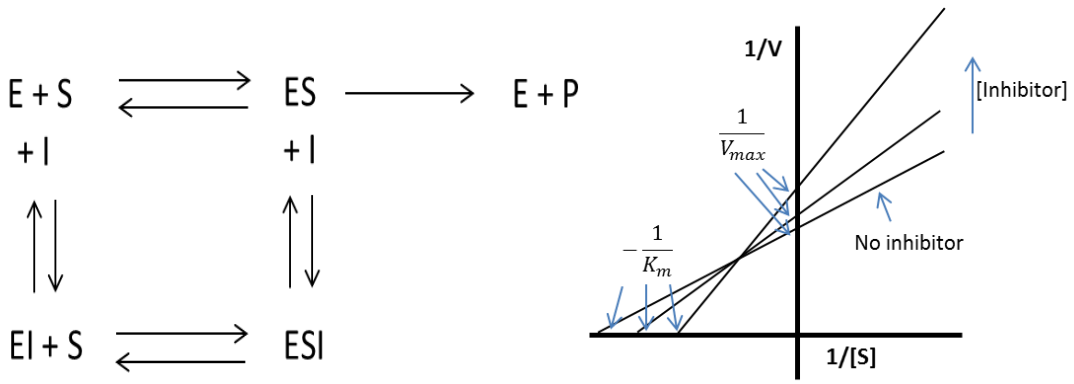


Figure 4.11. Non-competitive/ mixed inhibition. Inhibitor (I) binds at a separate site and may bind either to E (enzyme) or/and to ES (enzyme-substrate) complex (on the left). Lineweaver-Burk plot of mixed inhibition (on the right)

The inhibition mode can be determined by double-reciprocal plot (Lineweaver-Burk). The rate experiments are carried out in various substrate concentrations in the absence and in the presence inhibitor, with the enzyme concentration held constant. The results are plotted 1/reaction rate versus 1/substrate concentration.[257] In the presence of competitive inhibitor, increasing inhibition concentration results in a set of lines with a common intercept on the 1/V axis but with different slopes (V_{max} is unchanged). In the case of a mixed inhibitor, both intercept changes, resulting in changes in V_{max} and K_m (Figure 4.11).

The compound screening was carried out at a single compound concentration and the percentage inhibition was trend-setting. Therefore, further investigations were required to confirm the activity of the compounds. Five of the compounds: **p5**, **p6**, **p13**, **p15** and **p18** (Table 4.6) were chosen for the determination of inhibition constant (K_i) by the PCA. After the K_i determination, the fluorophore displacement in higher compound (**p5** and **p18**) concentrations was investigated (IC_{50}) by the LDA and finally the mode of inhibition was determined for **p5** and **p18**, in order to explain why the percentage of fluorophore displacement was low.

4.3.1.1 Determination of K_i

The PCA was used for the determination of the K_i values of the identified hits from the compound screening. Five of the compounds were selected based on their activity against LgtC. In order to investigate whether the compound specifically inhibited LgtC, the K_i values were determined in the presence and the absence of TX-100 evaluating the ability of the compounds to form aggregates and against β1,4-GalT. K_i values were determined as in the previous chapters for **p5**, **p6**, **p13**, **p15** and **p18**.

Firstly, the control test was carried out to ensure that the phosphatase was not inhibited by the pyrazolones prior to the activity experiments. A correlation between compound concentration and UDP formed in the phosphatase catalysed reaction was not observed (Figure 4.12), indicating that the compounds did not affect the efficiency of the PCA.

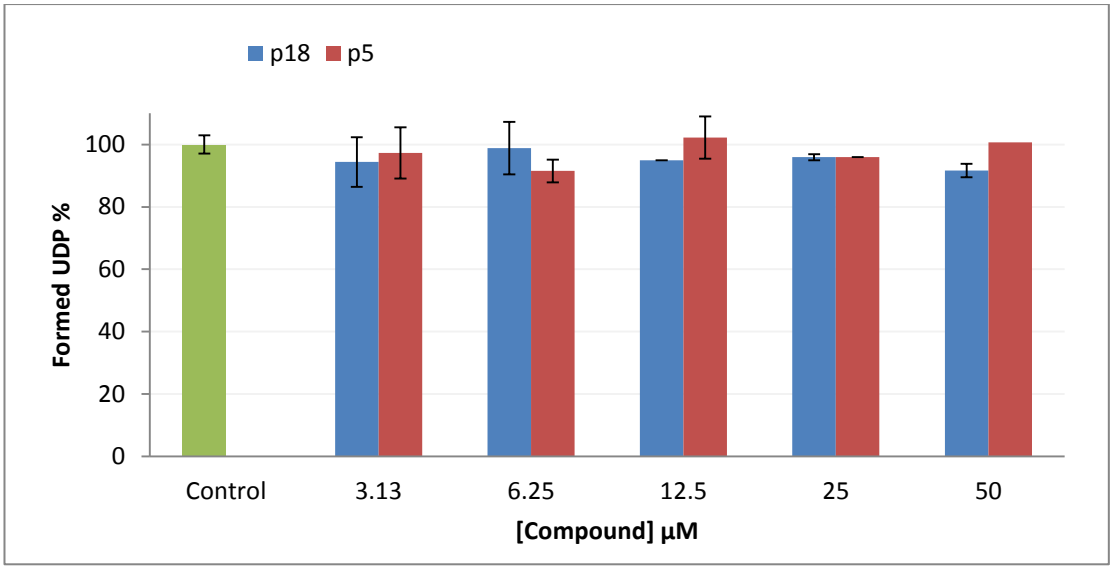


Figure 4.12. Phosphatase control by the PCA. The functionality of the phosphatase was not affected by UDP formation in various concentrations of **p5** or **p18**. Experiment was carried out at fixed concentrations of UDP (6.25 µM) and phosphatase (10 U/mL)

All tested pyrazolones, **p5**, **p6**, **p13**, **p15** and **p18**, inhibited LgtC in the range of 0.25 to 4.3 µM (Table 4.7). The K_i values remained constant in the presence of TX-100 indicating that the pyrazolones were not sensitive to the surfactant and specifically inhibited LgtC. In addition, the experiment performed against β1,4-GalT, confirmed the specific inhibition because the compounds did not show activity even at high compound concentration (Figure 4.13).

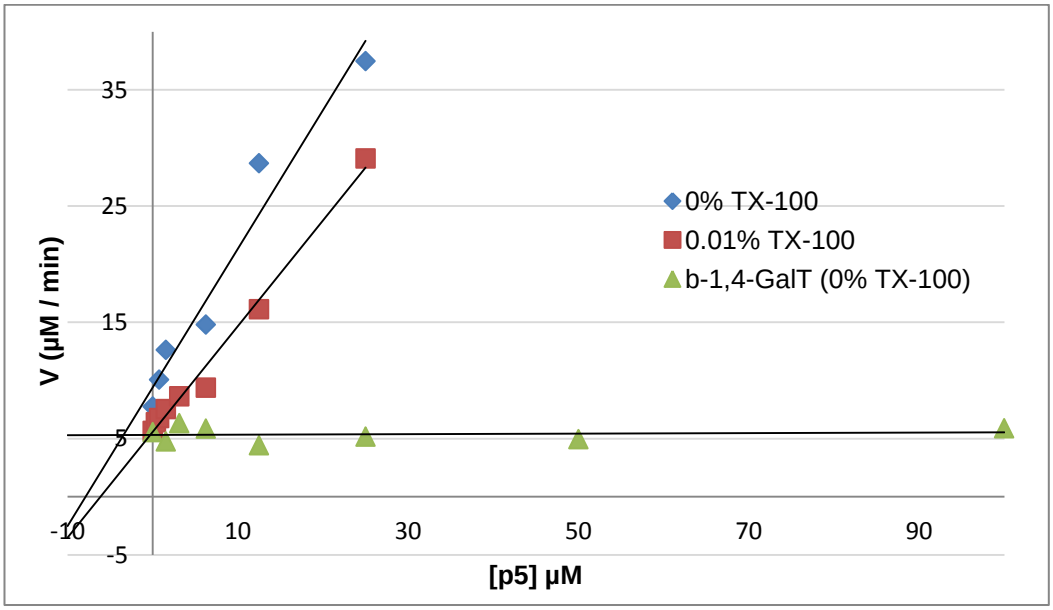


Figure 4.13. Dixon plot of reciprocal initial velocity (v) versus **p5** concentration to calculate the K_i value. K_i determination against LgtC in the absence and in the presence of TX-100, and against β1,4-GalT by the PCA

All the K_i experiments were performed without addition of Mn^{2+} to the assay buffer. As shown in Chapter 3.4.2 (Figure 3.22), Mn^{2+} increases the catalytic activity of LgtC, therefore, the K_i was performed for **p18** also in the excess of Mn^{2+}. K_i values were identical in the excess and without addition of Mn^{2+} (Table 4.7), indicating the inhibition was not affected by low Mn^{2+} concentration.

Table 4.7. Percentage inhibition by the PCA at 25 µM and K_i values against LgtC and β1,4-GalT in the presence and in the absence of TX-100

Compound	%INH	K_i (LgtC), µM		K_i (β1,4-GalT), µM
		0 %[a]	0.01 %[a]	
p5	94 ± 3	1.2	2.9	No inhibition observed
p6	86 ± 6	0.25	1.0	No inhibition observed
p13	79 ± 4	4.3	3.8	No inhibition observed
p15	88 ± 1	1.3	1.2	No inhibition observed
p18	86 ± 7	1.2 (1.1[b])	1.7 (1.0[b])	No inhibition observed

[a]TX-100 (v/v), [b]K_i values determined in the presence of 5.0 mM $MnCl_2$

The results strongly suggested that the pyrazolones specifically inhibited LgtC. Therefore, the structure-activity relationship was evaluated with **p5**, **p6**, **p13**, **p15** and **p18** based on the obtained K_i values. Based on the percentage inhibition results, it was suggested that carboxylic acid group is was not relevant to the inhibition and the similar K_i values of **p6** and **p15** confirmed this. In the absence of the carboxyl group at R_1 (**p13**, **p15** and **p18**) the compounds' activity against LgtC was not changed. The carboxylic acid group was demonstrated to be essential for the activity of the pyrazolones against WaaC,[250] due to the polar interaction with Arg on the acceptor binding site. However, Arg is not present on the acceptor binding site of LgtC[16] which possibly explains why the pyrazolones were active in the absence of the carboxylic acid group. The results suggested that the remaining part of pyrazolone structure interacts with LgtC. However, all the tested pyrazolones inhibited LgtC with the same concentration range and other conclusions regarding the binding could not be recognised.

4.3.1.2 Determination of IC_{50}

The K_i experiments by the PCA demonstrated that the pyrazolones were real LgtC inhibitors. The percentage fluorophore displacement by the LDA was significantly lower than in percentage inhibition by the PCA (Figure 4.10). Therefore, in order to investigate whether the pyrazolones were low affinity binders, the IC_{50} values were determined in the large concentration range. The experiment was performed in various **p5** and **p18** concentrations (up to 1.0 mM), and fixed LgtC and fluorophore concentrations. In addition, various pre-incubations were performed to investigate whether the observed inhibition is time-dependent:

- Incubation 1: **p5** or **p18** and fluorophore incubated for 10 minutes then LgtC added
- Incubation 2: **p5** or **p18** and LgtC incubated for 10 minutes then fluorophore added
- Incubation 3: LgtC and fluorophore incubated for 10 minutes then **p5** or **p18** added

The final reading was taken 20 minutes after the last component was added.

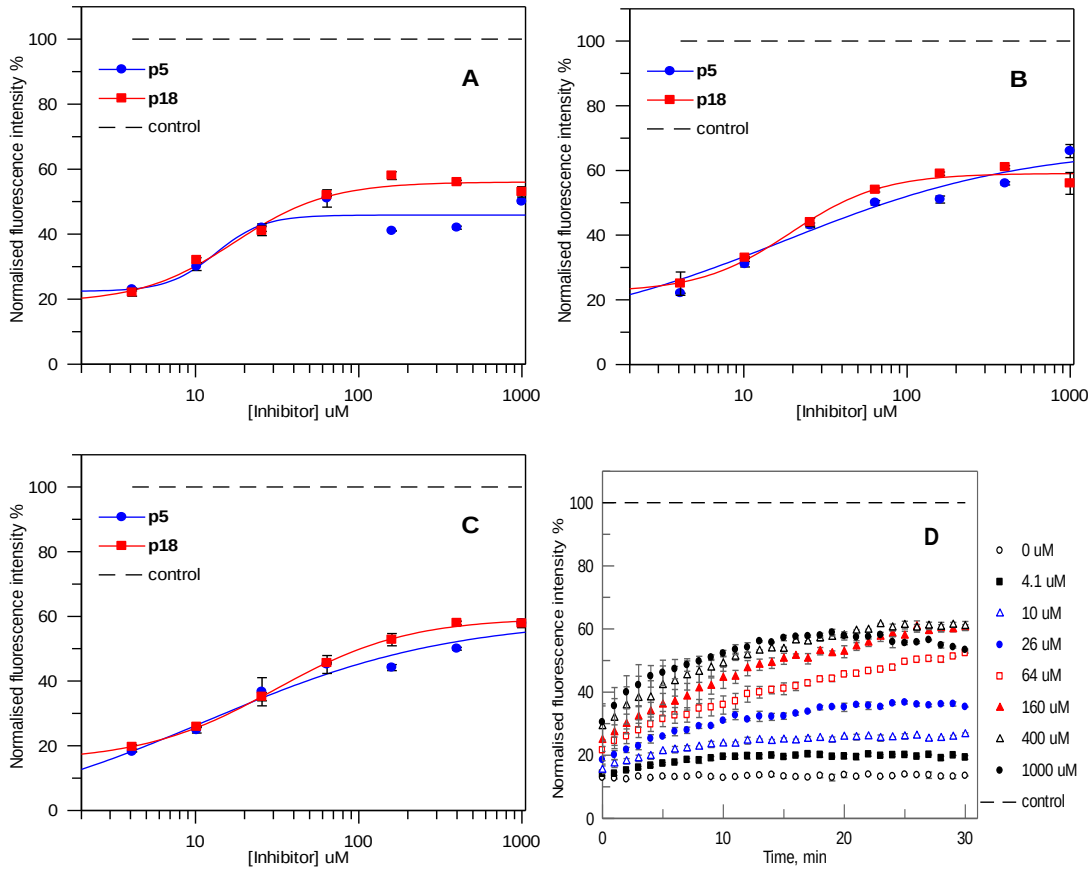


Figure 4.14. IC_{50} and pre-incubation study with **p5** and **p18** by the LDA in the presence of TX-100. A) Incubation 1, B) Incubation 2, C) Incubation 3, D) Incubation 3, the displacement of the fluorophore monitored in various concentration of **p18** over 30 minutes after **p18** was added

As illustrated in Figure 4.14, A, B and C, approximately 60 % of the fluorophore was displaced from the active site and the complete fluorescence recovery was not achieved at high concentration of **p5** or **p18**. A plateau was reached approximately at 70 µM of **p5** or **p18**. The fluorophore displacement by **p18** was followed for 30 minutes (Figure 4.14, D) and the fluorophore was first incubated with LgtC to ensure that fluorescence is fully quenched before **p18** was added. The curves demonstrated that **p18** slowly replaced the fluorophore reaching the plateau approximately after 15 minutes at high **p18** concentrations (> 64 µM).

The results indicated that **p5** and **p18** do not bind specifically to the donor binding site of LgtC. Possibly **p5** and **p18** binds to LgtC non-specifically causing a conformational change that releases the fluorophore from the active site. However, if **p5** and **p18** are mixed inhibitors then the partial fluorophore displacement could be explained because the compounds are capable of binding to free LgtC and the LgtC-fluorophore complex. If the pyrazolones were donor competitive inhibitors then the fluorescence would be restored completely. In the case of un-competitive inhibitors, the fluorescence would remain quenched as the inhibitor and the fluorophore share different binding sites. The LDA was unsuitable for the determination of the IC_{50} or the inhibition mode, however, the results strongly suggested that the pyrazolones were mixed inhibitors against the donor.

4.3.1.3 Determination of inhibition mode of pyrazolones

The inhibition of LgtC by the pyrazolones was confirmed with the K_i experiments. The percentage fluorophore displacement obtained by the LDA was significantly lower than the percentage inhibition obtained by the PCA (Figure 4.10) suggesting that the inhibition mode is more likely mixed than competitive inhibition. The partial fluorescence recovery was possibly due to the capability of pyrazolones to bind to substrate free (fluorophore) LgtC and the fluorophore-LgtC complex.

The mode of inhibition was first evaluated by performing Michaelis-Menten and Lineweaver-Burk plots in the absence and in the presence of **p5** or **p18** (at 2.0 and 5.0 µM). The same set of data was re-plotted (Lineweaver-Burk) and the type of inhibition was determined from the intersection of the slopes on the x- and y-axis.

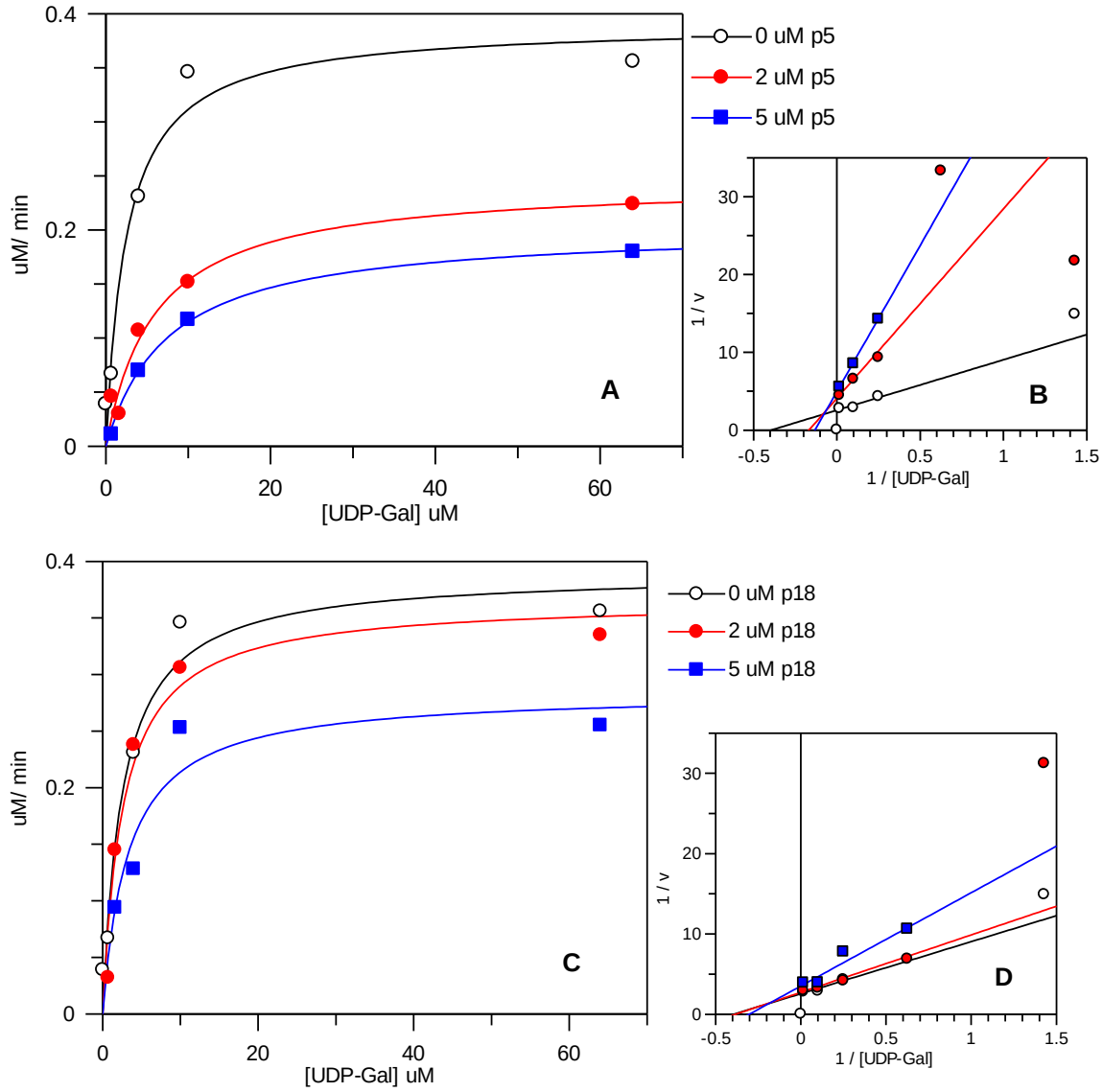


Figure 4.15. Michaelis-Menten and Lineweaver-Burk plot in the absence and in the presence of **p5** (A and B) or **p18** (C and D)

Both inhibitors followed Michaelis-Menten kinetics (Figure 4.15, A and C). The Lineweaver-Burk plot (Figure 4.15, B and D) illustrated that the straight lines had the same intercept above x-axis and changes in the x- and y-axis intercepts indicated mixed mode inhibition. Michaelis-Menten constants where determined at two different inhibition concentrations (Table 4.8). The K_m was 2.5 µM and V_{max} was 0.39 µM/min in the absence of inhibitor. K_m increased and V_{max} decreased with both **p5** and **p18** suggesting that less free enzyme is available, thus less substrate binding sites. The change in both, K_m and V_{max} values, supported the hypothesis that that **p5** and **p18** are mixed type of inhibitors against the donor.

Table 4.8. K_m (μM) and V_{max} (μM/ min) of UDP-Gal in the presence of **p5** or **p18** at 2.0 μM and 5.0 μM

[p5 or p18], μM	p5		p18	
	K_m	V_{max}	K_m	V_{max}
2.0	6.0 ± 1.7	0.25 ± 0.02	2.6 ± 0.8	0.37 ± 0.03
5.0	7.6 ± 0.8	0.20 ± 0.01	3.3 ± 1.6	0.28 ± 0.04

K_m= 2.5±1.0 μM and V_{max}= 0.39±0.04 μM/min in the absence of p5 **or** p18

Two graphical methods, Cornish-Bowden and Dixon plots,[258] were used in order to determine the K_i' for mixed inhibitors (**p5** and **p18**) and further investigate the mode of inhibition. By comparing these two plots in parallel for the same set of data, the mode of inhibition can be found from the pattern of the straight lines. Dixon plot (1/v versus inhibitor concentration) does not distinguish between competitive and mixed inhibitors. The intersection is found above x-axis in both cases and the inhibition constant is [I] = $-K_i$. The Cornish-Bowden plot ([S]/v versus inhibitor concentration) provides measure for enzyme-substrate-inhibitor (ESI) complex and, in case of competitive inhibitor the lines are parallel (no intersection) and for mixed inhibitors, the data for each substrate concentration fall on straight lines that intersect at [I] = $-K_i'$.

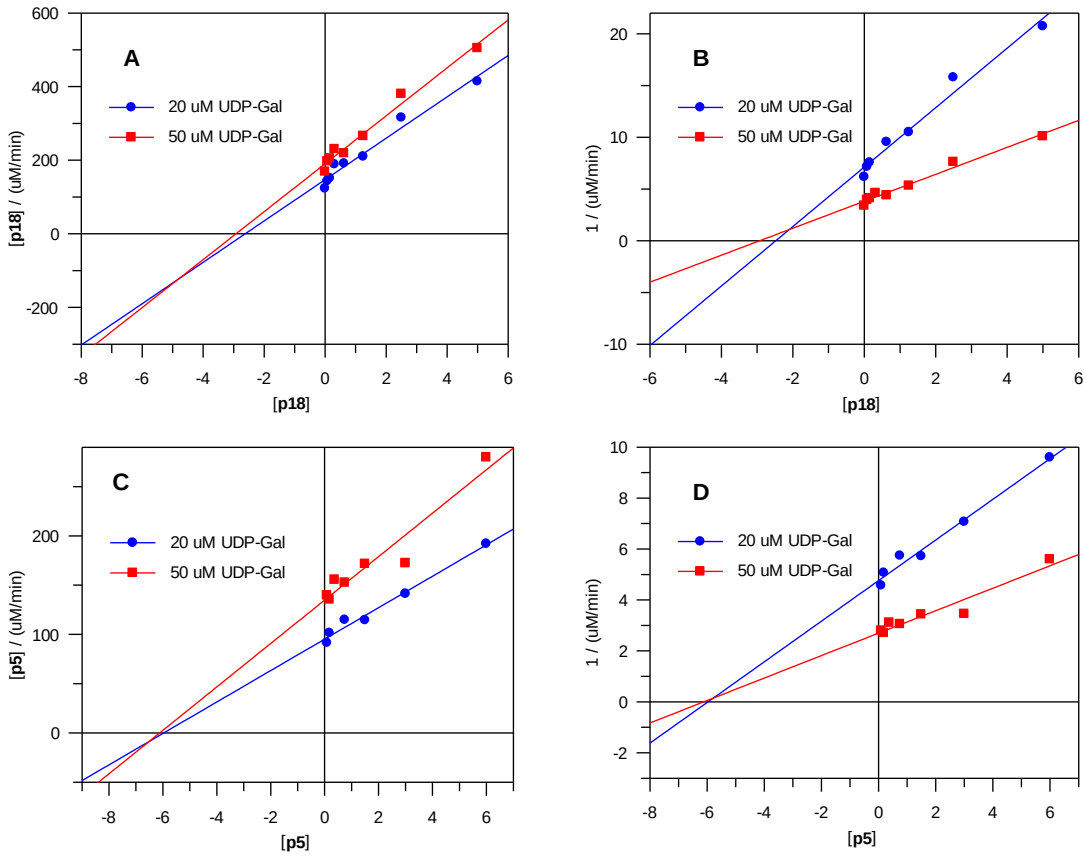


Figure 4.16. Cornish-Bowden plot (A and C) and Dixon plot (B and D) in various **p5** and **p18** concentrations. The intersection point on both plots suggests **p5** and **p18** are mixed mode inhibitors

The Cornish-Bowden and the Dixon plots were performed at 20 µM and 50 µM of UDP-Gal in various inhibitor concentrations. In the Cornish-Bowden plot, the intersection of two straight lines was below the x-axis and in the Dixon plot the lines intersected above the x-axis, indicating that **p5** and **p18** are mixed inhibitors.[258] The intersections of **p5** were near the x-axis which possibly indicates that **p5** is a non-competitive inhibitor instead of mixed inhibitor. However this was unlikely because all the results have indicated mixed inhibition thus far. Based on the Cornish-Bowden plots (intersection point is $-K_i'$) the calculated K_i' values were 5.8 µM and 6.5 µM for **p5** and **p18**, respectively.

The Michaelis-Menten kinetics indicated reversible, mixed inhibition and also the Cornish-Bowden plots indicated mixed inhibition. Therefore, **p5** and **p18** were mixed inhibitors and the re-calculated K_i values changed from 1.2 µM to 5.8 µM and 6.5 µM, respectively based on the preliminary results (Table 4.7). The mixed inhibition mode explains the results obtained with the LDA: the fluorescence was not restored in high **p5** or **p18** concentration, approximately 60 % of the fluorophore was displaced (Figure 4.14). Hypothetically, the fluorophore (**p5** and **p18**) can bind either to free LgtC or LgtC-fluorophore complex. If the LgtC-fluorophore-inhibitor complex is formed then part of the fluorophore remains on the donor binding site and cannot be displaced, i.e. LgtC-fluorophore-inhibitor complex, LgtC-fluorophore complex, free fluorophore and possibly LgtC-inhibitor complex are present at the same time.

4.4 Summary and conclusions

The LDA described in chapter 2 was used successfully for the screening of a large compound library. Nearly 400 compounds were screened in one day by one person which is in the HPLC-based assays due to their complex and time consuming protocols. The compound library screening included a significant number of pipetting steps which effected the accuracy of the results. The Z' factor (Figure 4.17) was approximately 0.8 or above for the six microplates prepared and after the sixth microplate, the Z' factor decreased. The decrease on the Z' is clearly due to the multiple pipetting steps and this error could be prevented if the screening was performed in an automated system.

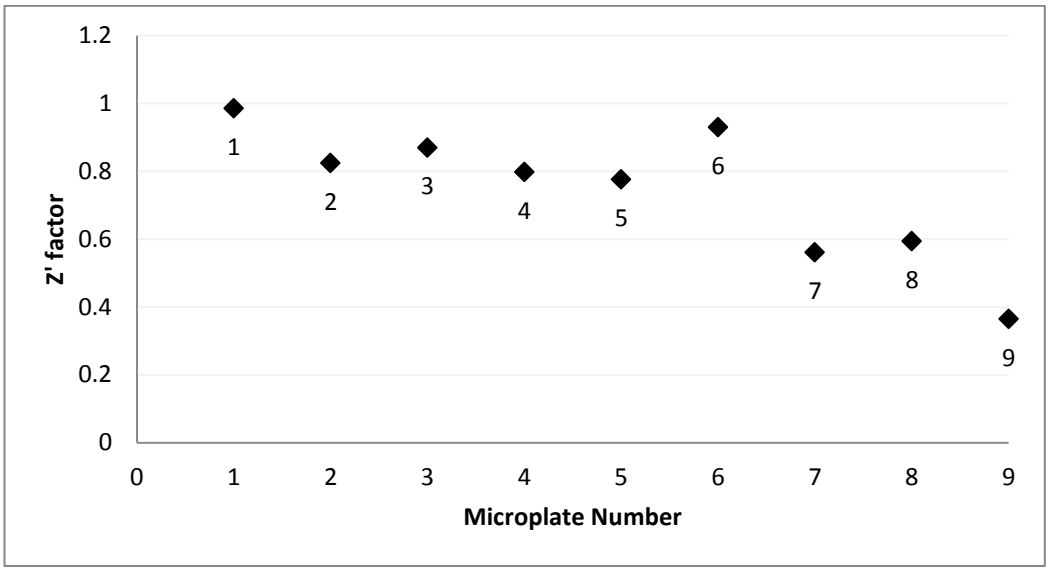


Figure 4.17. Z' factor reduced during the compound screening. Micro plate number 1 was prepared first and so forth to the last microplate number 9

Positive hits were identified from the DOS library screening. The surfactant sensitivity and non-specific inhibition investigations indicated that the hit compounds (**PP** analogues) were molecular aggregators. The particle size, determined by DLS confirmed that the hits compounds formed small particles in the assay conditions. These results supported the hypothesis that the small molecular aggregators non-specifically inhibited LgtC. Overall, these screening results showed that it is important to introduce a surfactant to the screening assay protocol to prevent the molecular aggregation.

Prior to the characterisation studies a small compound library of **PP** analogues was synthesised. The synthesis was affected by the poor solubility, specifically, the solubility of the impurities hampered the purification step. Two novel compounds were synthesised (**1e** and **1f**) and **1e** reached acceptable purity for further enzymological analysis. However, all the synthesised compounds were molecular aggregators. Generally, **PP** analogues were not drug-like molecules due their high molecular weight and high lipophilicity (LogP).

The focus-target library based on pyrazolones was screened and positive hits were found. The hits (**p5** and **p18**) inhibited LgtC at micro molar range and the further investigations indicated that they were mixed inhibitors. Therefore, these pyrazolones are not model drug-like molecules because they did not bind to the catalytic site of LgtC which may increase non-specific binding to other proteins. Generally, pyrazolones are classified as non-steroidal anti-

inflammatory drugs and they have been reported causing hypersensitivity.[259] For example, a well-known pain killer, metamizole, poses a risk inducing an agramulocytocis.[260] However, these pyrazolones could be potential targets for further development of specific LgtC inhibitors.

The target-focused library was screened parallel with the LDA and the PCA, in order to validate the LDA. The comparison of the results demonstrated that the percentage fluorophore displacement was lower with the LDA than the percentage inhibition with the PCA. Initially, a lack of shared control hampered the comparison. UDP was a suitable control in the LDA, however, not in the PCA because UDP is a recorded component and recognised by the phosphatase. A possibility to use the same control in both assays was investigated, however, either of the compounds tested were unsuitable for the purpose. $5CF_3$-UDP-Gal inhibited the phosphatase in the PCA and tunicamycin (a studied *N*-glycosylation inhibitor[252]) did not show activity against LgtC. The hits (pyrazolones) from the screening were used to investigate the difference between the results obtained with the LDA and the PCA. The hits were found to inhibit LgtC by a mixed inhibition mode which explained the low percentage fluorophore displacement. Hypothetically, the pyrazolones could bind to the free LgtC and also to the fluorophore-enzyme complex. Therefore, pyrazolones displace part of the fluorophore resulting in an increase in the fluorescence signal, however, because the pyrazolones can bind to the fluorophore-enzyme complex, part of the fluorophore remains on the donor binding site and fluorescence cannot be completely restored.

Overall, the compound library screenings showed that the LDA was simple and rapid to perform, therefore, it is suitable for HTS. Fluorescence interference of the tested compounds is low because the fluorescence intensities of the compounds and the fluorophore are compared before recording the final reading. The LDA is designed for the discovery of competitive inhibitors that displace the fluorophore from the donor binding site and it is unsuitable for the identification of un-competitive or mixed inhibitors. However, the LDA could be used for the characterisation inhibition mode in parallel with a biochemical assay because the fluorescence intensity could be fully recovered or fluorescence intensity would remain low for competitive and non-competitive inhibitors, respectively.

4.5 Experimental

A multi-mode BMG Labtech PolarStar microplate reader which has both, absorbance and fluorescence intensity detection modes was utilised for the compound screening and the characterisation experiments. All the compound screening and IC_{50} experiments by the LDA were carried out in black NUNC F96 MicroWell polystyrene plates or black Greiner 384 well micro polypropylene microplates. The fluorescence intensity was measured at 430 nm with excitation at 350 nm which were relevant to the fluorophore, 5FTUDP-Gal. The flashes per well was set to 10 and the gain was performed on blank well (fluorophore in buffer) and adjusted at 85 % on each experiment. All the compound screening and K_i experiments by the PCA were carried out in clear NUNC F96 MicroWell polystyrene plates and absorbance was recorded at 620 nm. Malachite reagents were prepared as described in experimental chapter 3.6.3. All reagent used in this chapter are listed in App. 2.

The compound screening was performed against LgtC and in the characterisation section, the hit compounds were tested against β1,4-GalT and the results were utilised as a control. The calculations were performed with excel and Michaelis-Menten kinetics always with GraFit version 7.0.3.

4.5.1 Diverse compound library screening

The compounds were received in 96 well microplates at 100 mM in 100 % DMSO. The compounds from the original plates were diluted to final concentration of 1.0 mM in 100 % DMSO to make “daughter plates”.

All the experiments were carried out in 13 mM HEPES (pH 7.0) buffer at final concentrations of LgtC (E), fluorophore (F) and DMSO at 10 µM, 0.2 µM and 10 % respectively (Table 4.9). The compounds were tested at 100 µM. Each microplate contained a negative control (F + E, low fluorescence), a positive control (F + E + 5.0 mM UDP, high fluorescence), a control (F + E + 50 µM UDP, 50 % fluorophore displacement) and a blank (F).

Table 4.9. Concentration of the components in the DOS library screening (96 well microplate)

Component	[Stock]	[Final]	Volume, µL
Inhibitor/ UDP control	1 mM[a]/ 0.5 and 50 mM[a]	100 µM /50 and 5000 µM	10
Fluorophore	1.0 µM	0.2 µM	20
LgtC	50 µM	10 µM	20
HEPES buffer	13 mM	13 mM	50

[a]In 100 % DMSO (v/v)

All the compounds and controls were run in duplicate (Figure 4.18). The compounds were incubated at 30 ˚C for 10 minutes without LgtC and first reading was taken. 10 minutes after the incubation LgtC or buffer to blank was added (total volume on the well 100 µL) and incubation was continued for 20 minutes before taking the final reading. The first and the final readings were used for calculation the percentage fluorophore displacement. The percentage fluorophore displacement was calculated as presented in chapter 4.1.1.

	1	2	3	4	5	6-12
A	Blank	Comp 1	Comp 1	Comp 9	Comp 9	Comp 16-40
B	Blank	Comp 2	Comp 2	Comp 10	Comp 10	Comp 17-41
C	Neg Ctrl	Comp 3	Comp 3	Comp 11	Comp 11	Comp 18-42
D	Neg Ctrl	Comp 4	Comp 4	Comp 12	Comp 12	Comp 1943
E	Pos Ctrl	Comp 5	Comp 5	Comp 13	Comp 13	Comp 20-44
F	Pos Ctrl	Comp 6	Comp 6	Comp 14	Comp 14	Comp 21-45
G	Ctrl	Comp 7	Comp 7	Comp 14	Comp 14	Comp 22-46
H	Ctrl	Comp 8	Comp 8	Comp 15	Comp 15	Comp 23-47

Figure 4.18. Layout of the microplate on the compound screening

4.5.2 Target-focused library screening

The target-focused library was screened in a similar fashion to the DOS library. The main difference was that TX-100 was included in the assay buffer (Table 4.10) and 384 well microplates were utilised instead of 96 well microplates. The reaction total volume was 20 µL and all compounds and controls were run in triplicate. The stock solutions (10 mM) of the tested compounds were prepared in 100 % DMSO and diluted to the desired concentration in 100 % DMSO. The compounds were tested at 25 µM. Otherwise, the assay set up was identical to the DOS library screening.

Table 4.10. Concentration of the components on the target-focused library screening (384 well microplate) by the LDA

Component	[Stock]	[Final]	Volume, µL
Inhibitor/ UDP control	250 µM[a] /0.25 and 50 mM[a]	25 µM /25 and 5000 µM	2
Fluorophore	1 µM	0.2 µM	4
LgtC	17 µM	5 µM	6
HEPES buffer	13 mM	13 mM	6
TX-100	0.1 % (v/v)	0.01 % (v/v)	2

[a]In 100 % DMSO (v/v)

The target-focused library was also screened with the PCA. All the experiments were carried out in 13 mM HEPES buffer (pH 7.0) at final concentrations of lactose (2.0 mM), CIP (10 U/mL), CEL (1.0 mg/mL), TX-100 (0.01 % (v/v)) and LgtC (60 nM), (Table 4.11). All the compounds were tested at 25 µM and the final DMSO concentration was 10 %.

Table 4.11. Concentration of the components on the target-focused library screening (96 well microplate) by the PCA

Component	[Stock]	[final]	V in well, µL
Lactose	10 mM	2 mM	30
CIP	100 U/mL	10 U/mL	15
CEL	10.0 mg/mL	1.0 mg/mL	15
TX-100	0.1 %[a]	0.01 %[a]	15
Inhibitor/ DMSO[b]	250 µM/ 100 %[a]	25 µM/ 10 %[a]	15
LgtC / Buffer[b, c]	300 nM/ -	60 nM/ -	30
UDP-Gal/ UDP[b]	150 µM/ 0- 62.5 µM	30 µM/ 0- 12.5 µM	30
Malachite regent A			30
Malachite regent B			30

[a](v/v), [b]in the UDP calibration/phosphatase control and negative control, [c]in the background hydrolysis

The master mixture (M) which contained lactose, CIP, CEL and TX-100 (total 75 µL) was added to each well followed by the compound addition. Pure 100 % DMSO was added to the negative control (no inhibition) and to the calibration/phosphatase control wells. LgtC was added all the wells except to the background and UDP calibration/phosphatase control wells, where HEPES buffer was added instead (Figure 4.19). Finally UDP-Gal or UDP (for the calibration) was added to initiate the reaction. The reaction mixture was incubated for 20 minutes at 30 °C and stopped by adding malachite reagent A. After malachite reagent B addition, the colour development was recorded for 30 minutes. The colour development was finished after approximately 20 minutes and the reading was used in percentage inhibition calculations.

	1	2	3	4	5	6	7	8
A								
B		**M**	**M**	M+G+E	M+G+E	M+G	M+G	
C		**M+U1**	**M+U1**	M+G+E+I1	M+G+E+I1	M+G+I1	M+G+I1	
D		**M+U2**	**M+U2**	M+G+E+I2	M+G+E+I2	M+G+I2	M+G+I2	
E		**M+U3**	**M+U3**	M+G+E+I3	M+G+E+I3	M+G+I3	M+G+I3	
F		**M+U4**	**M+U4**	M+G+E+I4	M+G+E+I4	M+G+I4	M+G+I4	
G		**M+U5**	**M+U5**	M+G+E+I5	M+G+E+I5	M+G+I5	M+G+I5	
H								

Figure 4.19. Layout of the microplate in compound screening by PCA. Columns 2-3; UDP calibration, columns 4-5; various compounds and negative control on row B, columns 6-7; background for each compound. Master mixture (M), UDP in various concentrations, 0, 0.78, 1.56, 3.13, 6.25 and 12.5 µM (U1-U5), UDP-Gal (G), LgtC (E) and various compounds (I)

The UDP calibration curve was used for the calculation the amount of UDP was formed in the reaction. Before calculating the percentage inhibition the background hydrolysis signal (no LgtC) was subtracted from the main signal. The percentage inhibition was calculated by normalising the amount of UDP formed to the negative control (100 % LgtC activity).

4.5.3 The hit characterisation studies

The LDA, the HPLC-based assay and the PCA were used for compound characterisation. The final DMSO concentration was 10 % (v/v) in all experiments. All inhibitor dilutions were in 100 % DMSO and pure DMSO was added to controls as a substitute. All the assay settings were the same as in the screening assays: the LDA was set as in chapter 4.5.1 and the PCA as in chapter 4.5.2.

LgtC was activated in 5 mM dithiothreitol (DTT) assay buffer for 30 minutes at 37 °C prior to use in the biochemical assays (HPLC and PCA). The final concentration of LgtC was fixed in each assay used in characterisation studies, however, the concentration varied depending on the assay, batch of LgtC and presence of Mn^{2+}. Concentration of LgtC was optimised on each assay prior the experiments by performing a calibration curve of LgtC and ensuring that the turnover of UDP-Gal to UDP was less than 10 %. Generally, the concentration of LgtC was 5 µM in LDA,

1.3 nM in HPLC, 60 nM and 25 nM in the absence and in the presence of Mn^{2+}, respectively in the PCA experiments.

4.5.3.1 Determination of IC_{50} and pre-incubation study by the LDA

IC_{50} experiments were performed at fixed concentration of LgtC and fluorophore (5FTUDP-Gal) and in the absence and in the presence of TX-100 (0.01 % (v/v)). The highest concentration of inhibitor (**p2_27** or **1c**) was 100 µM and diluted to 0.2 µM by 2 fold to obtain requisite data points to the curve. Negative (no UDP, low fluorescence intensity) and positive (5 mM UDP, high fluorescence intensity) controls were included on every microplate.

IC_{50} experiments were performed for **p5** and **p18** similarly as described above except the highest concentration of inhibitor was 1000 µM and diluted to 4.1 µM by 2.5 fold. The experiment was performed in the presence of TX-100 (0.01 % (v/v)).

The following pre-incubations were performed with (**p5** and **p18**):

- Incubation 1: Inhibitor and fluorophore incubated for 10 minutes then LgtC added
- Incubation 2: Inhibitor and LgtC incubated for 10 minutes then fluorophore added
- Incubation 3: LgtC and fluorophore incubated for 10 minutes then inhibitor added

The final reading was taken 20 minutes after the last component was added.

4.5.3.2 Determination of percentage inhibition by HPLC-based assay

HPLC settings were as in chapter 3.2.3. The experiment was carried out in eppendorf tubes containing UDP-Gal (10 µM), lactose (2.0 mM) and 50 µM inhibitor in 50 mM MOPS buffer (pH 7.0) including BSA (1.0 mg/mL) and $MnCl_2$ (20 mM). The reactions were initiated by adding activated LgtC (see experimental chapter 2.7.5 for activation). The total reaction volume was 100 µL. The reaction mixture was incubated at 37 °C for 10 minutes and stopped by placing the tubes in dry ice. The percentage inhibition was calculated by normalising the peak area of the inhibitor containing samples to the peak area of the negative control (maximum reaction velocity, no inhibitor added, 100 % LgtC activity).

4.5.3.3 Determination of activity of phosphatase in the presence of inhibitor

A phosphatase test of the PCA was carried out in in the absence of GT reaction, at fixed concentration of UDP (6.25 µM) and at 0, 6.25, 12.5, 25, 50 and 100 µM of required inhibitors. The reaction was initiated by adding UDP. (see Table 4.11 how the UDP calibration/ phosphatase control was performed, DMSO displaced with the compounds)

4.5.3.4 Determination of K_i by PCA

K_i was determined in various TX-100 concentrations (0, 0.001 and 0.01 %, (v/v)), in the absence and in the presence of Mn^{2+} (5.0 mM), and with LgtC and β1,4-GalT. All assay components were fixed except the inhibitor concentration (see Table 4.11). The initial velocity was determined at 0, 1.56, 3.13, 6.25, 12.5, 25, 50 and 100 µM of inhibitor. The measured data was converted to the reaction velocity and 1/v versus inhibitor concentration was plotted (Dixon plot). The K_i was determined at singe UDP-Gal concentration (30 µM assuming that the compounds were competitive inhibitors ($V = (Vmax\,[S])/(Km\,(1 + [I]/Ki\,) + [S]\,)$).

The experiments utilising β1,4-GalT were different from LgtC experiments. The experiments were carried out in 13 mM HEPES + 50 mM KCl buffer (pH 7.0), including GlcNAc (10 mM), CIP (10 U/mL), CEL (1.0 mg/mL), $MnCl_2$ (5.0 mM), β1,4-GalT (0.4 mU/mL) and UDP-Gal (18 µM). Otherwise, the assay was performed as PCA with LgtC (see Table 4.11 for the assay set up).

4.5.3.5 Determination of mode of inhibition by PCA

K_m was determined in the absence and in the presence of inhibitor, **p5** or **p18** (2.0 or 5.0 µM) and the same data was used in Lineweaver-Burk[261] plots (1/v vs. 1/[S], $1/V = ((Km + [S]))/((Vmax \times [S])\,) = (Km/Vmax\,) \times (1/[S]\, + 1/Vmax\,)$). The experiment was carried out in the presence of $MnCl_2$ (5.0 mM) and UDP-Gal at 0.7, 1.6, 4.0, 10, 26 and 64 µM (see Table 4.11 for the assay set up).

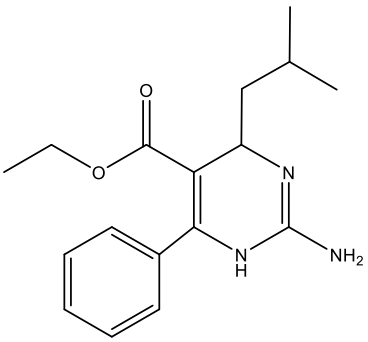
The Cornish-Bowden[258] (1/[S] vs. [I]) and the Dixon[262] (1/v vs. [I]) plots were performed in various inhibitor concentrations (0, 0.078, 0.156, 0.313, 0.625, 1.25, 2.5 and 5 µM) and at 20 µM and 50 µM of UDP-Gal. K_i was calculated from the intersection of the two straight lines (UDP-Gal 20 µM and 50 µM) in the Cornish-Bowden plot.

4.5.4 Synthesis of PP compounds

All reagents used in the compound synthesis were obtained commercially and used as received. All the reactions were followed by TLC in 6 % MeOH in $CHCl_3$ and the spots were visualized under UV light at 254/365 nm. The TLC plates were pre-coated slides of Silica Gel 60 F254 (Merck).

The synthesised compounds were identified and the preliminary purity was obtained by ^{1}H-NMR on a Bruker BioSpin GmbH at 400 MHz. The accurate purity was determined by HPLC (Chapter 4.5.5). ^{13}C-NMR spectra was recorded a Bruker BioSpin GmbH at 100 MHz and the accurate ESI-MS were obtained on a Finnigan MAT 900 XLT mass spectrometer at the EPSRC National Mass Spectrometry Service Centre, Swansea.

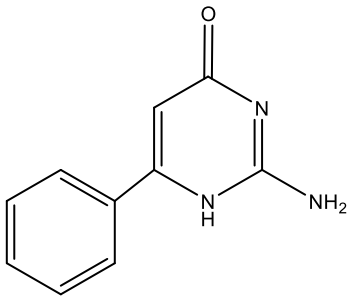
Ethyl 2-amino-4-isobutyl-6-phenyl-1,4-dihydropyrimidine-5-carboxylate (1a)



Ethylbenzoyl acetate (1.8 g, 9.583 mmol), guanidine hydrochloride (1.0 g, 10.5 mmol), isovaleraldehyde (1.02 mL, 9.5 mmol), sodium bicarbonate (2.93 g) and 100 mL DMF were loaded in a round bottom flask and refluxed at 70 °C for 16 hours. The reaction mixture including an orange precipitate was then poured over ice. After the ice had melted, the product was filtered and dried overnight in a desiccator. The yield of the freeze-dried solid

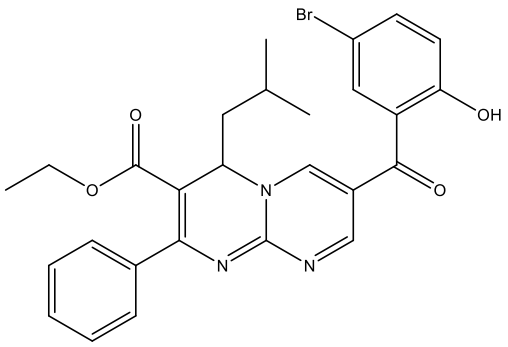
was 49 % (1.4 g). The compound was pure by ^{1}H-NMR. The purity was not obtained by HPLC because the compound was not visible by UV. ^{1}H-NMR (400 MHz; DMSO) δ= 7.38-7.10 (5H, m, aryl, C<u>H</u>), 6.18 (2H, br, s, N<u>H</u>$_2$), 4.26 (1H, dd, *J* 3.4, 9.3 Hz, N-C<u>H</u>), 3.87-3.67 (2H, m, C<u>H</u>$_2$CH$_3$), 1.85-1.71 (1H, m, C<u>H</u>(CH$_3$)$_2$), 1.49 (1H, ddd, *J* 4.1, 9.4, 13.4 Hz, CHC<u>H</u>$_2$CH), 1.17 (1H, ddd, *J* 3.5, 9.6, 13.2 Hz, CHC<u>H</u>$_2$CH), 0.93 (6H, dd, *J* 4.1, 6.5 Hz, (C<u>H</u>$_3$)$_2$), 0.85 (3H, t, *J* 7.1 Hz, CH$_2$C<u>H</u>$_3$), N<u>H</u> signal was not observed; ^{13}C-NMR (100 MHz; DMSO) δ= 166.0 (C), 155.5 (C), 144.9 (C), 142.4 (C), 128.2 (CH), 126.8 (CH), 126.7 (CH), 98.2 (C), 57.9 (CH$_2$), 47.5 (CH), 45.5 (CH$_2$), 23.9 (CH$_3$), 22.4 (CH), 21.5 (CH$_3$), 13.8 (CH$_3$); HRMS (ESI$^+$) *m/z* found 302.1864 [M-H]$^+$, $C_{17}H_{23}N_3O_2^+$ required 302.1863.

2-amino-6-phenylpyrimidin-4(1H)-one (2a)



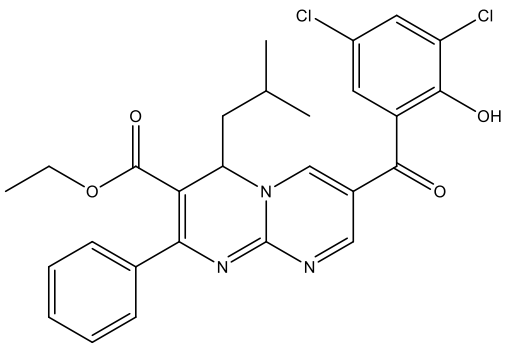
Ethylbenzoyl acetate (1.0 g, 5.208 mmol), guanidine carbonate (0.47 g, 5.222 mmol) and 100 mL EtOH were loaded in a round bottom flask and refluxed at 70 °C for 6 hours. The clear solution was reduced to dryness in vacuum, then EtOAc added and white precipitate formed. The product was then was reduced to dryness in vacuum and the solid material was washed with hot MeOH. The white precipitate was filtered and dried overnight in a desiccator. The yield of the freeze-dried solid was 36 % (346 mg). Pure by ^{1}H-NMR and HPLC (100%). ^{1}H-NMR (400 MHz; DMSO) δ= 10.8 (1H, br, s, N<u>H</u>), 7.99-7.92 (2H, m, aryl C<u>H</u>), 7.49-7-41 (3H, m, aryl C<u>H</u>), 6.64 (2H, br, s, N<u>H</u>$_2$), 6.13 (1H, s, C<u>H</u>); ^{13}C-NMR (100 MHz; DMSO) δ= 163.3 (C), 162.7 (C), 155.7 (C), 137.3 (C), 129.9 (CH), 128.4 (CH), 126.6 (CH), 97.4 (CH); HRMS (ESI$^+$) *m/z* found 188.0817 [M-H]$^+$, $C_{10}H_{10}N_3O^+$ required 188.0818.

Ethyl 7-(5-bromo-2-hydroxybenzoyl) -4-isobutyl-2-phenyl-4H-pyrimido [1,2-a]pyrimidine-3-carboxylate (1b)



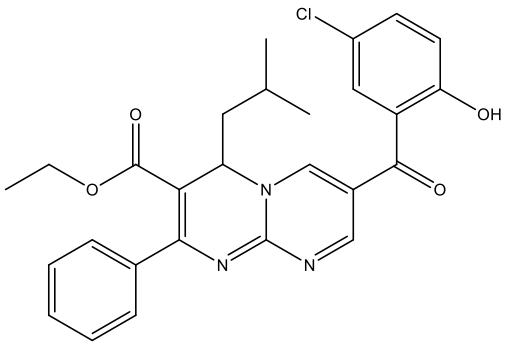
1a (50 mg, 0.166 mmol), 6-bromo-3-formylchromone (42 mg, 0.166 mmol) and 5 mL EtOH were loaded in a round bottom flask and refluxed at 70 °C for 5 hours. The precipitate was filtered hot and washed with hot EtOH. The precipitate was dried overnight in a desiccator. The yield of the freeze-dried solid was 46 % (41 mg) that was pure by ^{1}H-NMR and HPLC (96 %). ^{1}H-NMR (400 MHz; DMSO) δ= 10.67 (1H, br, s, O<u>H</u>), 7.87 (1H, d, *J* 2.5 Hz, aryl, C<u>H</u>), 7.75 (1H, dd, *J* 2.4, 8.7 Hz aryl, C<u>H</u>), 7.51-7.26 (7H, m, aryl, C<u>H</u>), 7.08 (1H, d, *J* 8.9 Hz, aryl, C<u>H</u>), 4.84 (1H, br, s, N-C<u>H</u>), 3.97-3.80 (2H, m, C<u>H</u>$_2$CH$_3$), 1.70 (3H, br, s, C<u>H</u>(CH$_3$)$_2$, CHC<u>H</u>$_2$CH), 1.08-0.89 (6H, m, (C<u>H</u>$_3$)$_2$), 0.85 (3H, t, *J* 7.2 Hz, CH$_2$C<u>H</u>$_3$); ^{13}C-NMR (100 MHz; DMSO) δ= 192.2 (C), 176.6 (C), 164.7 (C), 160.8 (C), 153.5 (C), 148.4 (CH), 141.7 (CH), 138.1 (CH), 134.7 (C), 129.1 (CH), 128.8 (CH), 128.7 (CH), 127.7 (CH), 121.0 (C), 113.6 (CH), 102.6 (C), 97.9 (C), 94.1 (C), 59.5 (CH$_2$), 49.5 (CH), 23.7 (CH), 23.4 (CH$_3$), 13.8 (CH$_3$); HRMS (ESI$^+$) *m/z* found 536.1170 [M-H]$^+$, $C_{27}H_{26}BrN_3O_4^+$ required 536.1179.

Ethyl 7-(3,5-dichloro-2-hydroxybenzoyl) -4-isobutyl-2-phenyl-4H-pyrimido [1,2-a]pyrimidine-3-carboxylate(1c)



1a (50 mg, 0.166 mmol), 6,8-dichloro-3-formylchromone (40 mg, 0.166 mmol) and 5 mL EtOH were loaded in a round bottom flask and refluxed at 70 °C for 5 hours. The precipitate was filtered hot and washed with hot EtOH twice. The precipitate was dried overnight in a desiccator. The yield of the freeze-dried solid was 39 % (34 mg) that was pure by ^{1}H-NMR and HPLC (95 %). ^{1}H-NMR (400 MHz; DMSO) δ= 10.52 (1H, br, s, O<u>H</u>), 7.95 (1H, s, aryl, C<u>H</u>), 7.72 (1H, s, aryl, C<u>H</u>), 7.66-7.23 (7H, m, aryl, C<u>H</u>), 4.95 (1H, br, s, N-C<u>H</u>), 4.02-3.80 (2H, m, C<u>H</u>$_2$CH$_3$), 1.72 (3H, br, s, C<u>H</u>(CH$_3$)$_2$, CHC<u>H</u>$_2$CH), 1.05-0.94 (6H, m, (C<u>H</u>$_3$)$_2$), 0.89 (3H, t, *J* 7.1 Hz, CH$_2$C<u>H</u>$_3$); ^{13}C-NMR (100 MHz; DMSO) δ= 186.7 (C), 177.0 (C), 170.9 (C), 158.0 (C), 154.7 (C), 147.5 (CH), 137.8 (CH), 134.7 (C), 129.0 (CH), 128.3 (C), 127.7 (CH), 123.7 (CH), 123.1 (C), 121.3 (CH), 101.1 (C), 103.8 (C), 100.8 (C), 81.9 (CH), 59.5 (CH$_2$), 59.1 (CH$_2$), 50.4 (CH), 23.2 (CH$_3$), 22.6 (CH$_3$), 22.0 (CH), 13.5 (CH$_3$); HRMS (ESI$^+$) *m/z* found 526.1287 [M-H]$^+$, $C_{27}H_{25}Cl_2N_3O_4^+$ required 526.1295.

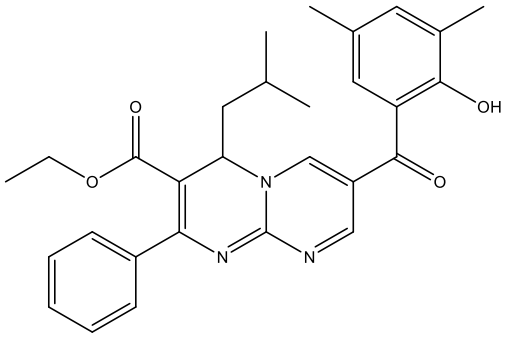
Ethyl 7-(5-chloro-2-hydroxybenzoyl) -4-isobutyl-2-phenyl-4H-pyrimido [1,2-a]pyrimidine-3-carboxylate (1d)



1a (50 mg, 0.166 mmol), 6-chloro-3-formylchromone (35 mg, 0.166 mmol) and 5 mL EtOH were loaded in a round bottom flask and refluxed at 70 °C for 5 hours. The precipitate was filtered hot and washed with hot EtOH twice. The precipitate was dried overnight in a desiccator. The yield of the freeze-dried solid was 55 % (45 mg) that was pure by ^{1}H-NMR and HPLC (99 %). ^{1}H-NMR (400 MHz; DMSO) δ= 10.67 (1H, br, s, O<u>H</u>), 7.74 (1H, d, *J* 2.6 Hz, aryl, C<u>H</u>), 7.64 (1H, dd, *J* 2.4, 8.6 Hz aryl, C<u>H</u>), 7.53-7.28 (7H, m, aryl, C<u>H</u>), 7.14 (1H, d, *J* 8.7 Hz, aryl, C<u>H</u>) 4.85 (1H, br, s, N-C<u>H</u>), 3.97-3.79 (2H, m, C<u>H</u>$_2$CH$_3$), 1.70 (3H, br, s, C<u>H</u>(CH$_3$)$_2$, CHC<u>H</u>$_2$CH), 1.10-0.91 (6H, m, (C<u>H</u>$_3$)$_2$), 0.85 (3H, t, *J*, 7.1 Hz, CH$_2$C<u>H</u>$_3$); ^{13}C-NMR (100 MHz; DMSO) δ= 193.8 (C), 177.5 (C), 164.9 (C), 160.4 (C), 153.8 (C), 148.1 (CH), 142.5 (CH), 137.4 (CH), 135.4 (C), 129.1 (CH), 128.8 (CH), 127.7 (CH), 125.6 (CH), 120.6 (C), 118.6 (CH), 103.9 (C), 93.1 (C), 83.2 (C), 59.5

(CH_2), 47.2 (CH), 23.4 (CH_3), 23.2 (CH), 13.4 (CH_3); HRMS (ESI^+) *m/z* found 492.1674 $[M-H]^+$, $C_{27}H_{26}ClN_3O_4^+$ required 492.1685.

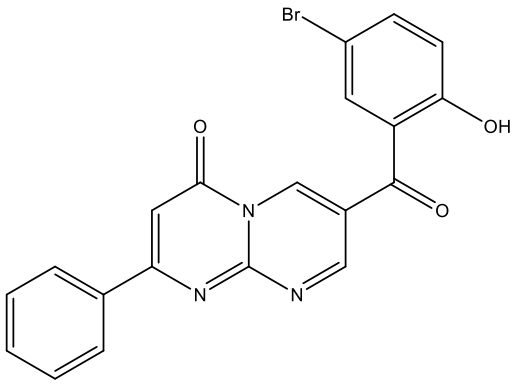
Ethyl 7-(2-hydroxy-3,5-dimethylbenzoyl) -4-isobutyl-2-phenyl-4H-pyrimido [1,2-a]pyrimidine-3-carboxylate (1e)



1a (50 mg, 0.166 mmol), 6,8-dimethyl-3-formylchromone (34 mg, 0.166 mmol) and 5 mL EtOH were loaded in a round bottom flask and refluxed at 70 °C for 5 hours. The precipitate was filtered hot and washed with hot EtOH twice. The precipitate was dried overnight in a desiccator. The yield of the freeze-dried solid was 51 % (41 mg) and pure by ^{1}H-NMR and HPLC (96 %). ^{1}H-NMR (400 MHz; DMSO) δ= 10.48 (1H, br, s, O<u>H</u>), 7.49-7.29 (9H, m, aryl, C<u>H</u>), 4.49 (1H, d, *J* 8.6 Hz N-C<u>H</u>), 4.01-3.79 (2H, m, C<u>H</u>$_2$CH$_3$), 1.84-1.63 (3H, m, C<u>H</u>(CH$_3$)$_2$, CHC<u>H</u>$_2$CH), 1.19-0.85 (15H, m, (C<u>H</u>$_3$)$_2$, CH$_2$C<u>H</u>$_3$); ^{13}C-NMR (100 MHz; DMSO) δ= 191.3 (C), 179.1 (C), 174.9 (C), 169.2 (C), 165.5 (C), 152.4 (CH), 149.2 (CH), 145.9 (C), 137.4 (CH), 135.3 (C), 131.0 (CH), 129.0 (CH), 127.5 (CH), 127.0 (C), 126.8 (CH), 124.3 (C), 116.6 (C), 105.6 (C), 59.4 (CH_2), 51.1 (CH), 41.5 (CH_2), 23.6 (CH_3), 23.3 (CH), 22.0 (CH_3), 22.1 (CH_3), 13.6 (CH_3); HRMS (ESI^+) *m/z* found 486.2380 $[M-H]^+$, $C_{29}H_{31}N_3O_4^+$ required 486.2387.

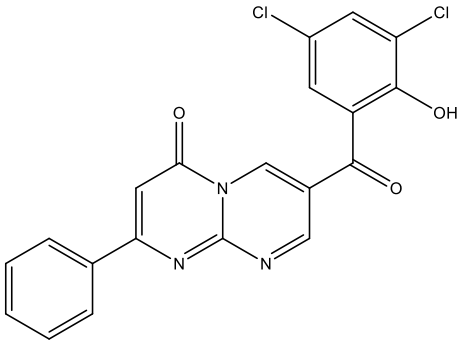
The following 4 compounds were not further analysed due to impurities and/or lack of material.

7-(5-bromo-2-hydroxybenzoyl) -2-phenyl-4H-pyrimido [1,2-a]pyrimidin-4-one (2b), $C_{20}H_{12}BrN_3O_3$, Mw 421.01



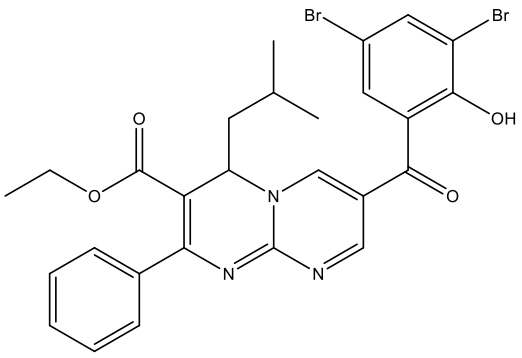
2a (50 mg, 0.267 mmol), 6-bromo-3-formylchromone (67 mg, 0.267 mmol) and 5 mL EtOH were loaded in a round bottom flask and refluxed at 70 °C for 5 hours. The precipitate was filtered hot and washed with hot MeOH twice. The precipitate was dried overnight in a desiccator. The compound was pure by ^{1}H-NMR and the yield of the freeze-dried solid was 2 % (2 mg).

7-(3,5-dichloro-2-hydroxybenzoyl) -2-phenyl-4H-pyrimido [1,2-a]pyrimidin-4-one (2c), $C_{20}H_{11}Cl_2N_3O_3$, Mw 411.02



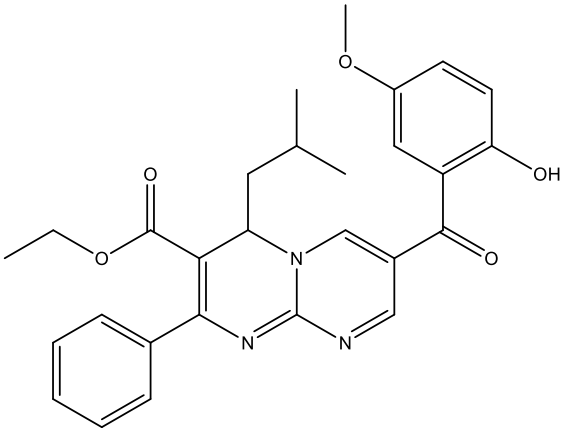
2a (80 mg, 0.428 mmol), 6,8-dichloro-3-formylchromone (104 mg, 0.428 mmol) and 8 mL EtOH were loaded in a round bottom flask and refluxed at 70 °C for 5 hours. The reaction was followed by TLC. This indicated that the reaction was incomplete. Another 1 eq of the chromone was added and the reaction was refluxed for a further 4 hours, resulting in no further conversion to the product. The precipitate was filtered hot and washed three times with hot EtOH and once with hot MeOH. The yield of the final freeze-dried solid was 26 % (45 mg) that was impure by ^{1}H-NMR and HPLC (24 %).

Ethyl 7-(3,5-dibromo-2-hydroxybenzoyl) -4-isobutyl-2-phenyl-4H-pyrimido [1,2-a]pyrimidine-3-carboxylate (1f), $C_{27}H_{25}Br_2N_3O_4$, Mw 613.02



1a (50 mg, 0.166 mmol), 6,8-dibromo-3-formylchromone (55 mg, 0.166 mmol) and 5 mL EtOH were loaded in a round bottom flask and refluxed at 70 °C for 5 hours. The precipitate was filtered hot and washed with hot EtOH twice. The precipitate was dried overnight in a desiccator. The yield of the final freeze-dried solid was 43 % (44 mg) that was impure by ^{1}H-NMR and HPLC (41 %).

Ethyl 7-(2-hydroxy-5-methoxybenzoyl) -4-isobutyl-2-phenyl-4H-pyrimido [1,2-a]pyrimidine-3-carboxylate(1g), $C_{28}H_{29}N_3O_5$, Mw 487.21



1a (50 mg, 0.166 mmol), 6-metho-3-formylchromone (34 mg, 0.166 mmol) and 5 mL EtOH were loaded in a round bottom flask and refluxed at 70 °C for 5 hours. White crystals were formed after keeping the reaction mixture in the refrigerator for four days. The crystals were identified as starting material by NMR. No reaction product formed.

4.5.5 Determination of λ_{max} and purity of the synthesised PP compounds

Prior to the purity measurements, the λ_{max} was determined for each compound by Varian Cary 50 Bio UV-VIS spectrophotometer with Cary Scan application. The compounds were analysed at 100 µM and diluted in water:ACN:DMSO (50:40:10, v/v). The measurements were carried out in quartz (1 cm) cuvettes and the compounds were scanned from 800 nm to 200 nm. Measured λ_{max} (Table 4.12) was utilised as a detection wavelength in the HPLC experiment.

Table 4.12. λ_{max}, HPLC gradient utilised and the purity of the synthesised compounds

Compound	λ_{max} (nm)	HPLC method	Purity (%)
1a	-	-	-
1b	420	G1	98
1c	420	G1	95
1d	380	G1	99
1e	380	G1	96
1f	380	G1	41
2a	310	G2	100
2c	310	G2	24

The purity of the synthesised compounds was determined by Perkin Elmer Series 200 HPLC, with Alltech Alltima C18 column (3 µm particle size, 15 cm x 4.6 mm), a DAD with desired detection wavelengths, a flow rate 1.0 mL/min and a column oven at 30 °C. The percentage purity of the compounds was calculated based on the peak area of the compound and the impurities. Two different gradient methods, varying water and ACN contents (Table 4.13) were developed and depending on the physical properties of the synthesised compounds, one of the two was utilised for analysis of purity.

Table 4.13. HPLC gradients utilised in the determination of the purity of the compounds

Time (min)	G1		G2	
	Water (%)	ACN (%)	Water (%)	ACN (%)
3	20	80	80	20
10	20-0	80-100	80-0	20-100
5	0	100	0	100
1	0-20	100-80	0-80	100-20
5	20	80	80	20

4.5.6 Determination of particle size by DLS

All compounds were analysed with Malvern Zetasizer Nano ZS instrument and particle size calculations were performed by Maplethorpe data analysis software. Measurements were carried out in disposable 12 mm square polystyrene cuvettes. The compounds were measured at a final concentration of 25 µM and compounds were diluted in 13 mM HEPES (pH 7.0) buffer and the DMSO concentration was 10 % (v/v), mimicking the in the LDA and the PCA assay conditions. Each particle size and intensity value represents three independent measurements. The measurements were carried out at 25 °C, the refractive index and viscosity of the diluent (10 : 90; water : DMSO) were set up to 1.346[247] and 1.06,[247] respectively. Refractive indexes were 1.580, 1.671, 1.642, 1.639, 1.632 and 1.614 for **1a**, **2a**, **1b**, **1c**, **1d** and **1e** respectively were determined by ACD/3D Viewer Freeware 12.01.

5 Summary and future work

In this project, assays for the discovery of novel GT inhibitors and for the investigation of GT activities have been developed. The LDA was successfully optimised and validated for the screening of compounds against LgtC. The PCA was demonstrated to be a time- and cost-effective method for the enzymological inhibitor characterisation compared e.g. with the HPLC-based method.

The LDA proved to be a robust and repeatable method during the assay optimisation. The enzyme and $MnCl_2$ concentrations had a significant influence on the IC_{50} values obtained. The optimised LDA was performed without $MnCl_2$ addition because the stock solution of LgtC contained $MnCl_2$ and a high concentration of $MnCl_2$ reduced the sensitivity of the assay. If the LDA is applied to other GTs than LgtC then it is important to optimise the $MnCl_2$ concentration with the desired enzyme. The surfactant, TX-100, had a positive influence for the assay performance: in the absence of TX-100, the fluorescence signal was not completely restored in the ligand displacement experiments and after introducing TX-100, the fluorescence was completely restored, resulting in an improvement in the assay signal window (S/B ratio) and in the reproducibility (Z' factor) of the LDA. The practicability of the LDA was demonstrated by screening a library of 393 diverse compounds in one day. Although, the LDA is operationally simple to perform, the adaptation to the automated system would be advantageous because errors during a manual operation could be avoided and screening a large compound library would be more practical. The compound library screening also demonstrated that the LDA can be only used for the identification of competitive inhibitors and that it is important to include a surfactant in the assay protocol in early stages to prevent possible false positives. Currently, the LDA is suitable for searching potential inhibitors against LgtC, GTB and α1,3-GalT. After this study, it is unknown how the fluorescent UDP-Gal analogue (5FTUDP-Gal) binds to the catalytic site of these GTs and which interactions upon binding results in the fluorescence quenching. Hypothetically, a conformational change of the fluorophore occurs upon binding and causes the drop in fluorescence signal. A crystal structure of the fluorophore in the complex with LgtC or α1,3-GalT would provide useful information about the possible conformational changes of the fluorophore, which amino acids are involved, as well as how many fluorophore molecules bind in relation to the protein.

A novel fluorescence-based glycosidase coupled assay was designed by coupling the GT reaction with fluorescent unnatural acceptor (Rgal) and glycosidase reaction. The major drawback during the development of this assay was the possible activity of LgtC on α1,4-linked galactose in the absence of the natural acceptor, lactose. The results obtained with the HPLC-based method indicated that UDP-Gal was consumed rapidly compared with Rgal. The further investigations with LC-MS/MS showed the formation of multiple polygalactosylated products, Rgal-gal and Rgal-gal-gal, based on their molecular weights and retention times. The LC-MS/MS experiments were performed in the absence of Rgal-gal and Rgal-gal-gal standards. Standards would provide important information regarding accurate retention times of the products and the correct parent and product ions which assists on compound identification. In conclusion, these results indicated that it is essential to design the LgtC assay using lactose as the acceptor preventing the α1,4-linked galactose serving as an acceptor. Alternatively, if an unnatural fluorescent acceptor is required, then designing an acceptor with binding affinity that is higher than the α1,4-linked galactose is essential for the assay performance. A limitation of the assays where a fluorescent acceptor is used is the narrow linear concentration range of the fluorescent molecule.

The PCA was successfully applied to LgtC and the robustness of the method was demonstrated with enzymological studies. The PCA was also demonstrated as having a high tolerance to DMSO and being an operationally simple method which was advantageous for inhibitor characterisation. A disadvantage of the PCA is that the microplates fill up with controls. In addition to the background hydrolysis and phosphatase controls, the wells on the outer edge are left empty which means that the K_m experiment only can be performed at seven different UDP-Gal concentrations (+ background hydrolysis) in triplicate. This is only a concern when the assay is used for the determination of the mode of inhibition and the K_m is performed in various inhibitor concentrations. To avoid performing the assay multiple times to obtain the desired results it would be advantageous to apply the PCA to 384 well microplates.

Two compound libraries were screened and two classes of hits were identified. One class proved to be molecular aggregators that resulted in non-specific inhibition and the other class was a real inhibitor against LgtC. The further investigations demonstrated that the compounds (pyrazolones) inhibited LgtC at micromolar range, however, they were mixed inhibitors. The results obtained by the ligand displacement assay suggested that the pyrazolones bind to the donor binding site because the fluorescence was partly recovered. Initially, the pyrazolones

were designed as acceptor analogue inhibitors of WaaC.[250] Therefore, hypothetically, if the pyrazolones bind to acceptor binding site then the fluorophore remain on the donor binding site and fluorescence cannot be recovered (Figure 5.1, A). This hypothesis could be investigated by performing the assay in the presence of acceptor (Figure 5.1, B). A mixture of acceptor, fluorophore and LgtC would be incubated first, to ensure that the fluorescence signal is quenched. Then incubation in the presence of inhibitor would result a high fluorescence signal if the pyrazolones binds to the acceptor binding site and to the donor binding site. However, it is also possible that the pyrazolones bind to a location other than to the active site.

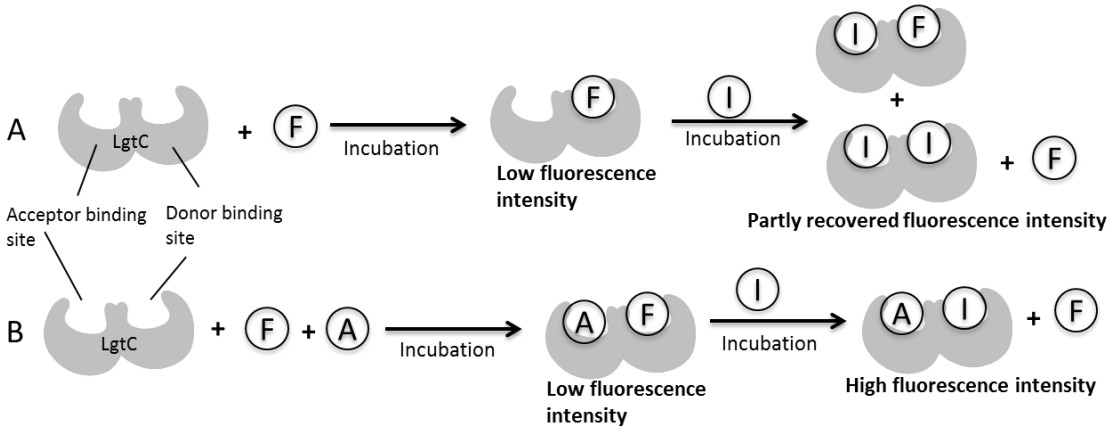


Figure 5.1. A) Possible outcome of the LDA when the fluorophore is displaced with pyrazolones. B) The principle of the fluorophore displacement in the presence of acceptor in order to investigate if pyrazolones bind to the acceptor binding site. Abbreviations: A; acceptor, F; fluorophore, I; inhibitor

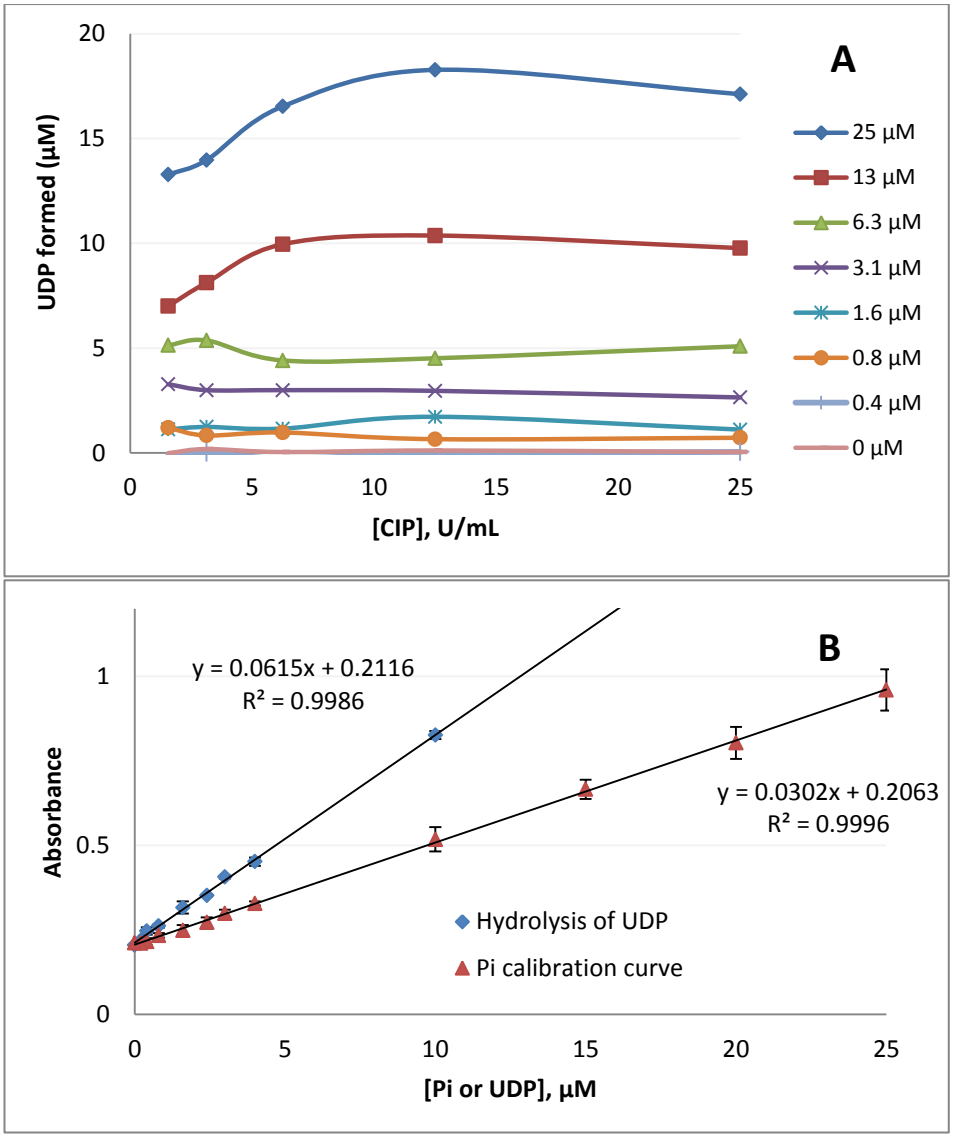
If the pyrazolones bind to the donor and to the acceptor binding sites, or if one pyrazolone molecule binds to the donor-acceptor binding site then they could be modified to a specific competitive inhibitors. However, further developments require a detailed docking study or a three-dimensional structure of the LgtC-inhibitor complex.

Appendix

1. Phosphatase coupled assay development with β1,4-GalT

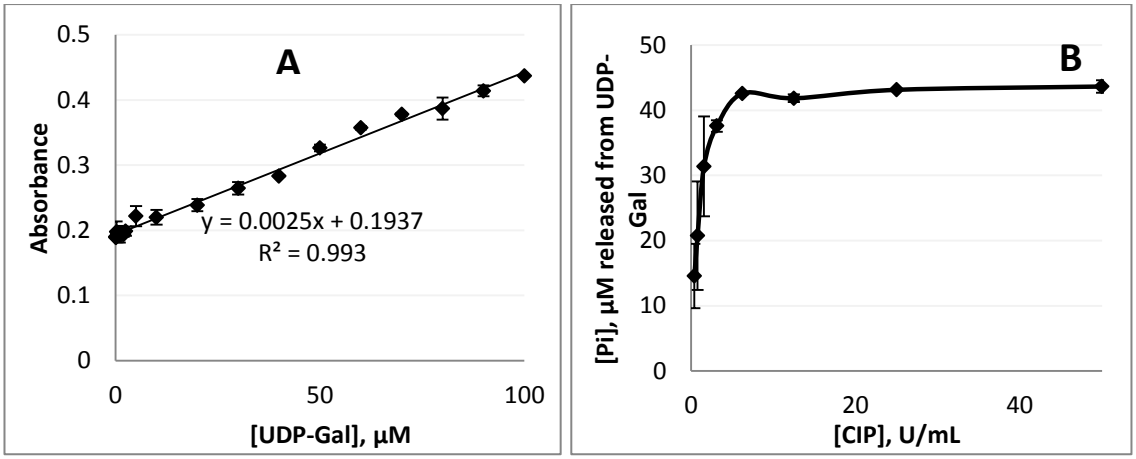
During the course of this project, a novel phosphatase coupled assay was reported[85]. Difficulties to reproduce this assay in house were encountered. Dr Evitt carried out this optimisation study with β1,4-GalT. Because the assay conditions used in this thesis are not yet published they are presented here.

Dr Evitt discovered an alternative phosphatase which cleavages the phosphate group specifically only from UDP and which was financially more viable than the phosphatase used in the reported PCA:[85] a calf intestinal alkaline phosphatase (CIP) that catalyses the cleavage of terminal phosphate from the 5'phophorylated ends of DNA or RNA and is widely used in molecular biology.[263] The optimum CIP concentration was examined by varying CIP and UDP concentrations (App. 1, A) and formed phosphate was detected by malachite reagents.[85] The plateau was reached approximately at 10 U/mL of CIP at 12.5 and 25 µM of UDP whilst at lower UDP concentrations less CIP was required to release phosphate. In order to determine how many phosphates CIP at 10 U/mL hydrolyse from UDP, a standard curve of UDP and inorganic phosphate at the same concentrations were compared. At the tested range 10 U/mL CIP hydrolysed 2 equivalents of inorganic phosphate from UDP (App. 1, B).



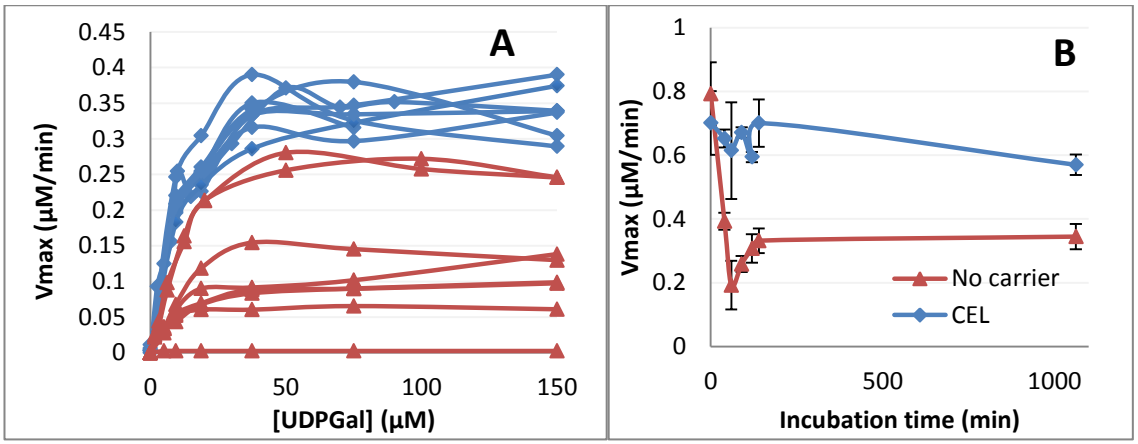
App. 1. A) Optimisation of CIP concentration in various UDP concentrations. B) Formation of phosphate by hydrolysis of UDP vs. phosphate standard curve

Dr Evitt investigated if UDP-Gal behaves as a substrate for CIP by performing a standard curve of UDP-Gal. The results indicated that if CIP releases phosphate from UDP-Gal it is from an impurity or hydrolysis product, UDP (App. 2, A). This was further investigated at fixed concentration of UDP-Gal with increasing concentration of CIP (App. 2, B). The results clearly demonstrated that UDP-Gal was not a substrate of CIP because 43 µM of phosphate at most was released from 600 µM of UDP-Gal, indicating that UDP-Gal contained approximately 7 % phosphate impurity. The phosphate impurity was then subtracted from the main signal by including a control in each UDP-Gal concentration in the absence of acceptor. In addition, the same control was utilised for subtracting the enzymatic/chemical hydrolysis of UDP-Gal.



App. 2. A) Formation of phosphate by hydrolysis of UDP-Gal. B) Increasing concentration of CIP at 600 µM UDP-Gal

β1,4-GalT was used as a model enzyme to develop suitable conditions for enzymatic kinetic studies. Dr Evitt applied the assay to determine K_m of the donor substrate UDP-Gal. However, under the current assay conditions V_{max} vary significantly (App. 3, A, red curves). The results were reasoned that under the assay conditions, β1,4-GalT may be susceptible to attach to plastic surfaces (*i.e*. microplate wells, pipette tips etc.) and consequently to affect the active enzyme concentration which can be significant as β1,4-GalT is present in low concentrations in the assay. The active concentration of the enzyme and the length of exposure to the plastic can vary, therefore, affecting the assay results. This problem is often improved by the adding of a carrier protein such as BSA to the assay buffer. However, Dr Evitt discovered that BSA significantly prevents the colour development of the malachite reagents thus an alternative carrier protein was required.



App. 3. A) Replicates of Michaelis-Menten plots for β1,4-GalT varying UDP-Gal substrate concentration. B) Time course data for β1,4-GalT in the presence and in the absence of hen egg white-cell lysozyme, CEL. Red curves: experiments conducted in the absence of carrier protein CEL. Blue curves: identical experiments conducted in the presence of 1 mg/mL CEL

Various compounds such as single amino acids (*e.g.* Cys, Leu), surfactants (*e.g.* TX-100) and a selection of proteins (*e.g.* BSA) were tested as potential carrier proteins. Hen egg white-cell lysozyme (CEL) was a potential option because the colour development of the malachite reagents was not hindered in its presence. The further investigations showed significant improvement in the reproducibility of V_{max} (App. 3, A, blue curves) in the presence of CEL and in the time-course experiment, the activity of β1,4-GalT dropped by half over 200 minutes in the absence of CEL (App. 3, B). The results clearly showed that CEL acted as a carrier protein and β1,4-GalT preserved full activity in the presence of CEL. Dr Evitt improved the PCA significantly by using CIP and CEL in the assay set up and during the optimisation the assay demonstrated to be a useful tool for GT activities.

The optimised phosphatase coupled assay conditions are presented in App. 4. All the β1,4-GalT reactions were carried out in 13 mM HEPES + 50 mM KCl (pH 7.0) buffer, including $MnCl_2$ (5 mM), DMSO (10% (v/v)), CEL (1 mg/ml), CIP (10 U/ml), β1,4-GalT (0.4 mU/mL), GlcNAc (10 mM). UDP-Gal was added as a starter to the β1,4-GalT reactions and UDP to the calibration curve/phosphatase control. The reactions were incubated for 20 minutes at 30 °C and stopped by adding malachite reagent A and mixed thoroughly, and then malachite reagent B was added and again mixed thoroughly. The colour development was monitored for 30 minutes.

App. 4. Phosphatase coupled biochemical assay set up

Assay step	Component	[Stock]	[reaction]	Volume in well, µL
1	GlcNAc/ Buffer	50 mM	10 mM	30
	CIP	100 U/ mL	10 U/ mL	15
	$MnCl_2$	50 mM	5.0 mM	15
	CEL	10 mg/ mL	1.0 mg/ mL	15
	DMSO	100 %	10 %[a]	15
2	β1,4-GalT/ Buffer	5x	1x	30
3	UDP-Gal/ UDP	5x	1x	30
4	**20 min incubation at 30 °C**			
5	Malachite reagent A	5x	1x	30
6	Malachite reagent B	5x	1x	30
7	**colour development monitored at 620 nm for 30 minutes**			

[a] v/v

2. List of reagents

Reagent	Purity	Supplier
ACN	HPLC grade	Fisher Scientific
Ammonium Acetate	98.7 %	VWR
Ammonium molybdate (para) tetrahydrate	99%	Alfa Assar
Ampicillin	-	MP Biomedicals
6-bromo-3-formylchomone	99 %	Sigma
BSA	>98 %	Sigma
Casamino acids	-	MP Biomedicals
CEL (Hen egg white lysozyme)	-	Merck
6-chloro-3-formylchomone	97 %	Acros
CIP (calf intestinal phosphatase)	-	New England Biolabs
Dialysis tubing	-	Sigma
6,8-dibromo-3-formylchomone	99 %	Sigma
6,8-dichloro-3-formylchomone	97 %	Acros
6,8-dimethyl-3-formylchomone	-	SantaCruz Biotechnology
DMF (dimethylformamide)	≥99.5 %	Sigma
DTT	>99 %	Sigma-Aldrich
Enolpyruvate kinase	-	Sigma
EtOH	99.9 %	VWR
galactosidase (β-D-Galactoside galactohydrolase, Lactase from Aspergillus oryzae)	-	Sigma
GlcNAc	99 %	Fluka
Glucose	-	Lab M Limited
guanidine carbonate	-	Fluka
Glycerol	-	MP Biomedicals
HCl	37 %	Sigma-Aldrich
HEPES	≥99.5 %	Sigma-Aldrich
Imidazole	99 %	Fisher Scientific
IPTG (isopropyl-β-D-thiogalactopyranoside)	-	Bioline Ltd.
isovaleraldehyde		Sigma
KCl	99 %	BDH
Lactate dehydrogenase	-	Calbiochem
Lactose (α-D-Lactose monohydrate)	97 %	Alfa Aesar
M9 Minimal Salts, 5X	-	Sigma
Malachite green oxalate	-	Alfa Aesar
$MgCl_2$	≥98 %	Sigma
6-metho-3-formylchomone	99 %	Sigma

Reagent	Purity	Supplier
MeOH	HPLC grade	Fisher Scientific
$MnCl_2 \cdot 4H_2O$	≥99 %	Acros
MOPS	99 %	Acros
NaCl	Extra pure	Acros
NADH (β-nicotinamide adenine dinucleotide, reduced form)	≥95 %	Sigma
NuPAGE 4-12% Bis-Tris 12well gel	-	Invitrogen
NuPAGE antioxidant solution	-	Invitrogen
NuPAGE LDS sample buffer (4X)	-	Invitrogen
PAGE ruler prestained protein marker	-	Fermentas
Nutrient Agar	-	Fisher Scientific
Phosphoenolpyruvate	99 %	Sigma
Potassiumphosphate dibasic, (K_2HPO_4)	≥99.0 %	Fluka
Potassiumphosphate monobasic, (KH_2PO_4)	≥99.5 %	Fluka
PVA (Polyvinyl alcohol)	99%	Alfa Aesar
Rgal (Resorufin-β-D-galactopyranoside)	≥95 %	Fluka
Resorufin	95 %	Sigma-Aldrich
SimplyBlue safestain	-	Invitrogen
TBAHS (tetrabutylammonium hydrogen sulphate)	≥99.0 %	Sigma-Aldrich
TRIS	99.8 %	Acros
TritonX-100	Bioextra	Sigma
Tryptone	-	Fisher Scientific
UMP	98 %	Acros
UDP	≥99.5 %	Sigma
UDP-Gal	97.1 %	CalbioChem
uridine	>99 %	Sigma
Vitamin-B1 (thiamine)	-	MP Biomedicals
Yeast extract	-	Lab M Limited

References

1. J. Rini, J. Esko and A. Varki, *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, New york, 2009.
2. L. L. Kiessling and R. A. Splain, *Annu Rev Biochem*, 2010, **79**, 619-653.
3. C. A. Weijers, M. C. Franssen and G. M. Visser, *Biotechnol Adv*, 2008, **26**, 436-456.
4. C. Breton, S. Fournel-Gigleux and M. M. Palcic, *Curr Opin Struct Biol*, 2012, **22**, 540-549.
5. L. L. Lairson, B. Henrissat, G. J. Davies and S. G. Withers, *Annu Rev Biochem*, 2008, **77**, 521-555.
6. R. Caputto, L. F. Leloir, C. E. Cardini and A. C. Paladini, *J Biol Chem*, 1950, **184**, 333-350.
7. L. F. Leloir, *Science*, 1971, **172**, 1299-1303.
8. G. J. Seifert, *Curr Opin Plant Biol*, 2004, **7**, 277-284.
9. G. Baisch, R. Ohrlein and A. Katopodis, *Bioorg Med Chem Lett*, 1997, **7**, 2431-2434.
10. G. Baisch, R. Ohrlein, M. Streiff and B. Ernst, *Bioorg Med Chem Lett*, 1996, **6**, 755-758.
11. B. R. Griffith and J. S. Thorson, *Nat Chem Biol*, 2006, **2**, 659-660.
12. L. L. Lairson, A. G. Watts, W. W. Wakarchuk and S. G. Withers, *Nat Chem Biol*, 2006, **2**, 724-728.
13. L. L. Lairson, W. W. Wakarchuk and S. G. Withers, *Chem Commun (Camb)*, 2007, 365-367.
14. C. P. Chiu, A. G. Watts, L. L. Lairson, M. Gilbert, et al., *Nat Struct Mol Biol*, 2004, **11**, 163-170.
15. L. N. Gastinel, C. Bignon, A. K. Misra, O. Hindsgaul, et al., *EMBO J*, 2001, **20**, 638-649.
16. K. Persson, H. D. Ly, M. Dieckelmann, W. W. Wakarchuk, et al., *Nat Struct Biol*, 2001, **8**, 166-175.
17. B. L. Cantarel, P. M. Coutinho, C. Rancurel, T. Bernard, et al., *Nucleic Acids Res*, 2009, **37**, D233-238.
18. CarbohydrateActive Enzymes server http://www.cazy.org/, 2013.
19. Y. Liu, X. Li, H. Xu and H. Qiao, 2008, 78-81.
20. S. J. Charnock and G. J. Davies, *Biochemistry*, 1999, **38**, 6380-6385.
21. B. Ramakrishnan, P. V. Balaji and P. K. Qasba, *J Mol Biol*, 2002, **318**, 491-502.
22. A. Vrielink, W. Ruger, H. P. Driessen and P. S. Freemont, *EMBO J*, 1994, **13**, 3413-3422.
23. S. Morera, L. Lariviere, J. Kurzeck, U. Aschke-Sonnenborn, et al., *J Mol Biol*, 2001, **311**, 569-577.
24. Y. Yuan, D. Barrett, Y. Zhang, D. Kahne, et al., *Proc Natl Acad Sci U S A*, 2007, **104**, 5348-5353.
25. M. Igura, N. Maita, J. Kamishikiryo, M. Yamada, et al., *EMBO J*, 2008, **27**, 234–243.
26. C. Breton, L. Snajdrova, C. Jeanneau, J. Koca, et al., *Glycobiology*, 2006, **16**, 29R-37R.
27. I. Tvaroska, *Carbohydr Res*, 2004, **339**, 1007-1014.
28. H. Gomez, I. Polyak, W. Thiel, J. M. Lluch, et al., *J Am Chem Soc*, 2012, **134**, 4743-4752.
29. K. Ohtsubo and J. D. Marth, *Cell*, 2006, **126**, 855-867.
30. M. Abu-Qarn, J. Eichler and N. Sharon, *Curr Opin Struct Biol*, 2008, **18**, 544-550.
31. E. Weerapana and B. Imperiali, *Glycobiology*, 2006, **16**, 91R-101R.
32. H. Nothaft and C. M. Szymanski, *Nat Rev Microbiol*, 2010, **8**, 765-778.
33. D. H. Dube and C. R. Bertozzi, *Nat Rev Drug Discov*, 2005, **4**, 477-488.
34. A. Alavi and J. S. Axford, *Rheumatology (Oxford)*, 2008, **47**, 760-770.
35. H. Ghazarian, B. Idoni and S. B. Oppenheimer, *Acta Histochem*, 2011, **113**, 236-247.
36. T. F. Orntoft and E. M. Vestergaard, *Electrophoresis*, 1999, **20**, 362-371.
37. J. Sun, J. Thurin, H. S. Cooper, P. Wang, et al., *Proc Natl Acad Sci*, 1995, **92**, 5724-5728.
38. J. Renkonen, T. Paavonen and R. Renkonen, *Int J Cancer*, 1997, **74**, 296-300.

39. W. Vollmer, D. Blanot and M. A. de Pedro, *FEMS Microbiol Rev*, 2008, **32**, 149-167.
40. M. Jacques, *Trends Microbiol*, 1996, **4**, 408-409.
41. N. M. Young, J. R. Brisson, J. Kelly, D. C. Watson, et al., *J Biol Chem*, 2002, **277**, 42530-42539.
42. C. R. Raetz and C. Whitfield, *Annu Rev Biochem*, 2002, **71**, 635-700.
43. S. Piek and C. M. Kahler, *Front Cell Infect Microbiol*, 2012, **2**, 162.
44. M. P. Jennings, Y. N. Srikhanta, E. R. Moxon, M. Kramer, et al., *Microbiology*, 1999, **145**, 3013-3021.
45. P. Zhu, M. J. Klutch, M. C. Bash, R. S. Tsang, et al., *Microbiology*, 2002, **148**, 1833-1844.
46. W. W. Wakarchuk, M. Gilbert, A. Martin, Y. Wu, et al., *Eur J Biochem*, 1998, **254**, 626-633.
47. L. Y. Lin, B. Rakic, C. P. Chiu, E. Lameignere, et al., *J Biol Chem*, 2011, **286**, 37237-37248.
48. W. Wakarchuk, A. Martin, M. P. Jennings, E. R. Moxon, et al., *J Biol Chem*, 1996, **271**, 19166-19173.
49. M. P. Jennings, D. W. Hood, I. R. A. Peak, M. Virji, et al., *Mol Microbiol*, 1995, **18**, 729-740.
50. Y. L. Tzeng and D. S. Stephens, *Microbes Infect*, 2000, **2**, 687-700.
51. T. Kajimoto and M. Node, *Synthesis*, 2009, **19**, 3179-3210.
52. K. Descroix, T. Pesnot, Y. Yoshimura, S. S. Gehrke, et al., *J Med Chem*, 2012, **55**, 2015-2024.
53. J. Paszkowska, K. Kral, T. Bieg, U. Nawrot, et al., *Molecules*, 2013, **18**, 8018-8027.
54. K. Schaefer, J. Albers, N. Sindhuwinata, T. Peters, et al., *Chembiochem*, 2012, **13**, 443-450.
55. Y. Hu, J. S. Helm, L. Chen, C. Ginsberg, et al., *Chem Biol*, 2004, **11**, 703-711.
56. B. J. Gross, J. G. Swoboda and S. Walker, *J Am Chem Soc*, 2008, **130**, 440-441.
57. C. D. Rillahan, S. J. Brown, A. C. Register, H. Rosen, et al., *Angew Chem Int Ed Engl*, 2011, **50**, 12534-12537.
58. R. Roychoudhury and N. L. Pohl, *Curr Opin Chem Biol*, 2010, **14**, 168-173.
59. J. R. Brown, B. E. Crawford and J. D. Esko, *Crit Rev Biochem Mol Biol*, 2007, **42**, 481-515.
60. K. C. Seo, Y. G. Kwon, D. H. Kim, I. S. Jang, et al., *Chem Commun (Camb)*, 2009, 1733-1735.
61. Y. Gao, C. Lazar, W. A. Szarek and I. Brockhausen, *Glycoconj J*, 2010, **27**, 673-684.
62. J. R. Brown, F. Yang, A. Sinha, B. Ramakrishnan, et al., *J Biol Chem*, 2009, **284**, 4952-4959.
63. G. K. Wagner and T. Pesnot, *Chembiochem*, 2010, **11**, 1939-1949.
64. T. D. Butters, *Curr Opin Chem Biol*, 2007, **11**, 412-418.
65. T. D. Butters, R. A. Dwek and F. M. Platt, *Glycobiology*, 2005, **15**, 43R-52R.
66. M. M. Palcic and S. Keiko, *Trends Glycosci Glyc*, 2001, **13**, 361-370.
67. M. M. Palcic, L. D. Heerze, M. Pierce and O. Hindsgaul, *Glycoconj J*, 1988, **5**, 49-63.
68. O. von Ahsen, U. Voigtmann, M. Klotz, N. Nifantiev, et al., *Anal Biochem*, 2008, **372**, 96-105.
69. F. Goulard, M. Diouris, E. Deslandes and J. Y. Floc'h, *Eur J Phycol*, 1999, **34**, 21-25.
70. I. Meynial, V. Paquet and D. Combes, *Anal Chem*, 1995, **67**, 1627-1631.
71. Y. Hayashi, Y. Horibata, K. Sakaguchi, N. Okino, et al., *Anal Biochem*, 2005, **345**, 181-186.
72. K. R. Anumula, *Glycobiology*, 2012, **22**, 912-917.
73. Y. Kanie, A. Kirsch, O. Kanie and C. H. Wong, *Anal Biochem*, 1998, **263**, 240-245.
74. A. Monegal, R. Pinyol and A. Planas, *Anal Biochem*, 2005, **346**, 115-123.
75. K. B. Lee, U. R. Desai, M. M. Palcic, O. Hindsgaul, et al., *Anal Biochem*, 1992, **205**, 108-114.

76. K. Sujino, T. Uchiyama, O. Hindsgaul, N. O. L. Seto, et al., *JACS*, 2000, **122**, 1261-1269.
77. J. Wu, S. Takayama, C. H. Wong and G. Siuzdak, *Chem Biol*, 1997, **4**, 653–657.
78. A. J. Norris, J. P. Whitelegge, K. F. Faull and T. Toyokuni, *Biochemistry*, 2001, **40**, 3774-3779.
79. A. J. Norris, J. P. Whitelegge, K. F. Faull and T. Toyokuni, *Anal Chem*, 2001, **73**, 6024-6029.
80. M. Yang, M. Brazier, R. Edwards and B. G. Davis, *Chembiochem*, 2005, **6**, 346-357.
81. C. Deng and R. R. Chen, *Anal Biochem*, 2004, **330**, 219-226.
82. M. Persson and M. M. Palcic, *Anal Biochem*, 2008, **378**, 1-7.
83. H. C. Hang, C. Yu, K. G. Ten Hagen, E. Tian, et al., *Chem Biol*, 2004, **11**, 337-345.
84. S. Gosselin, M. Alhussaini, M. B. Streiff, K. Takabayashi, et al., *Anal Biochem*, 1994, **220**, 92-97.
85. Z. L. Wu, C. M. Ethen, B. Prather, M. Machacek, et al., *Glycobiology*, 2011, **21**, 727-733.
86. A. J. Pope, U. M. Haupts and K. J. Moore, *Drug Discov Today*, 1999, **4**, 350-362.
87. J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Springer, New York, USA, 2006.
88. P. Gribbon and A. Sewing, *Drug Discov Today*, 2003, **8**, 1035-1043.
89. J. S. Helm, Y. Hu, L. Chen, B. Gross, et al., *J Am Chem Soc*, 2003, **125**, 11168-11169.
90. B. J. Gross, B. C. Kraybill and S. Walker, *J Am Chem Soc*, 2005, **127**, 14588-14589.
91. M. Solero-Higgin, E. E. Carlson, J. H. Philips and L. L. Kiessling, *J Am Chem Soc*, 2004, **126**, 10532-10533.
92. J. Wongkongkatep, Y. Miyahara, A. Ojida and I. Hamachi, *Angew Chem Int Ed Engl*, 2006, **45**, 665-668.
93. A. Chandrasekaran, K. Deng, C. Y. Koh, T. Takasuka, et al., *Chem Commun (Camb)*, 2013, **49**, 5441-5443.
94. P. Bojarova and V. Kren, *Trends Biotechnol*, 2009, **27**, 199-209.
95. D. G. Naumoff, *Biochemistry (Mosc)*, 2011, **76**, 622-635.
96. B. Henrissat, G. Sulzenbacher and Y. Bourne, *Curr Opin Struct Biol*, 2008, **18**, 527-533.
97. G. Davies and B. Henrissat, *Structure*, 1995, **3**, 853-859.
98. C. Bertoldo and G. Antranikian, *Meth Enzymol*, 2001, **330**, 269-289.
99. Z. D. Shi, O. Motabar, E. Goldin, K. Liu, et al., *Anal Bioanal Chem*, 2009, **394**, 1903-1909.
100. C. A. Mapes and C. C. Sweeley, *Biochem Biophys Res Commun*, 1973, **53**, 1317–1324.
101. F. Yagi, A. E. Eckhardt and I. J. Goldstein, *Arch Biochem Biophys*, 1990, **280**, 61-67.
102. T. Pesnot, PhD Thesis, Novel Sugar-Nucleotides for the Investigation of Glycosyltransferases, University of East Anglia, 2009.
103. L. Yan, D. F. Smith and R. D. Cummings, *Anal Biochem*, 1994, **223**, 111-118.
104. T. Pesnot, M. M. Palcic and G. K. Wagner, *Chembiochem*, 2010, **11**, 1392-1398.
105. S. H. Shin and K. S. Kim, *EOP*, 2013, **13**, 2189-2206.
106. R. Borrow, *Trop Med Int Health*, 2012, **17**, 1478-1491.
107. G. Boccadifuoco, J. Donnelly, D. Medini, M. Giuliani, et al., presented in part at the Meningitis and Septicaemia in Children and Adults 2011, Poster V36, London, UK, 2011.
108. P. H. Chan, S. Weissbach, M. Okon, S. G. Withers, et al., *Biochemistry*, 2012, **51**, 8278-8292.
109. H. D. Ly, B. Lougheed, W. W. Wakarchuk and S. G. Withers, *Biochemistry*, 2002, **41**, 5075-5085.
110. D. W. Hood, M. E. Deadman, M. P. Jennings, M. Bisercic, et al., *Proc Natl Acad Sci*, 1996, **93**, 11121-11125.
111. R. Griffin, C. D. Bayliss, M. A. Herbert, A. D. Cox, et al., *Infect Immun*, 2005, **73**, 7022-7026.

112. N. A. Spahich, D. W. Hood, E. R. Moxon and J. W. St Geme, 3rd, *J Bacteriol*, 2012, **194**, 1815-1822.
113. A. L. Erwin, S. Allen, D. K. Ho, P. J. Bonthuis, et al., *Infect Immun*, 2006, **74**, 6226-6235.
114. D. K. Ho, S. Ram, K. L. Nelson, P. J. Bonthuis, et al., *J Immunol*, 2007, **178**, 1002-1012.
115. D. Tsao, K. L. Nelson, D. Kim and A. L. Smith, *Microbes Infect*, 2012, **14**, 509-516.
116. B. Ramakrishnan and P. K. Qasba, *J Mol Biol*, 2001, **310**, 205-218.
117. B. Ramakrishnan, E. Boeggeman, V. Ramasamy and P. K. Qasba, *Curr Opin Struct Biol*, 2004, **14**, 593-600.
118. E. G. Berger and J. Rohrer, *Biochimie*, 2003, **85**, 261-274.
119. B. Ramakrishnan, E. Boeggeman and P. K. Qasba, *J Biol Chem*, 2012, **287**, 28666-28674.
120. L. N. Gastinel, C. Cambillau and Y. Bourne, *EMBO*, 1999, **18**, 3546-3557.
121. P. K. Qasba, B. Ramakrishnan and E. Boeggeman, *Curr Drug Targets*, 2008, **9**, 292-309.
122. V. Ramasamy, B. Ramakrishnan, E. Boeggeman and P. K. Qasba, *J Mol Biol*, 2003, **331**, 1065-1076.
123. H. J. Hathaway and B. D. Shur, *J Cell Biol*, 1992, **117**, 369-382.
124. C. Rodeheffer and B. D. Shur, *Biochim Biophys Acta*, 2002, **1573**, 258-270.
125. Y. Han, X. Zhou, Y. Ji, A. Shen, et al., *Cell Immunol*, 2010, **262**, 11-17.
126. Q. Lu and B. D. Shur, *Development*, 1997, **124**, 4121-4131.
127. M. Asano and K. Furukawa, *EMBO*, 1997, **16**, 1850-1857.
128. H. J. Choi, T. W. Chung, C. H. Kim, H. S. Jeong, et al., *Biochem Biophys Res Commun*, 2012, **426**, 620-625.
129. S. Villegas-Comonfort, N. Serna-Marquez, O. Galindo-Hernandez, N. Navarro-Tito, et al., *J Cell Biochem*, 2012, **113**, 3330-3341.
130. F. Zhu, F. Shen, Y. Fan, Y. Xie, et al., *Glycoconj J*, 2012, **29**, 347-356.
131. X. Sun, T. Savidge and H. Feng, *Toxins (Basel)*, 2010, **2**, 1848-1880.
132. D. E. Voth and J. D. Ballard, *Clin Microbiol Rev*, 2005, **18**, 247-263.
133. T. Dingle, S. Wee, G. L. Mulvey, A. Greco, et al., *Glycobiology*, 2008, **18**, 698-706.
134. D. J. Reinert, T. Jank, K. Aktories and G. E. Schulz, *J Mol Biol*, 2005, **351**, 973-981.
135. T. Jank, T. Giesemann and K. Aktories, *Glycobiology*, 2007, **17**, 15R-22R.
136. G. Pfeifer, J. Schirmer, J. Leemhuis, C. Busch, et al., *J Biol Chem*, 2003, **278**, 44535-44541.
137. S. Genisyuerek, P. Papatheodorou, G. Guttenberg, R. Schubert, et al., *Mol Microbiol*, 2011, **79**, 1643-1654.
138. J. Reineke, S. Tenzer, M. Rupnik, A. Koschinski, et al., *Nature*, 2007, **446**, 415-419.
139. A. Shen, P. J. Lupardus, M. M. Gersch, A. W. Puri, et al., *Nat Struct Mol Biol*, 2011, **18**, 364-371.
140. K. J. Choi, S. Grass, S. Paek, J. W. St Geme, 3rd, et al., *PLoS One*, 2010, **5**, e15888.
141. F. Kawai, S. Grass, Y. Kim, K. J. Choi, et al., *J Biol Chem*, 2011, **286**, 38546-38557.
142. F. Schwarz, Y. Y. Fan, M. Schubert and M. Aebi, *J Biol Chem*, 2011, **286**, 35267-35274.
143. S. Grass, C. F. Lichti, R. R. Townsend, J. Gross, et al., *PLoS One*, 2010, **6**, 1-9.
144. V. S. Stoll and J. S. Blanchard, *Meth Enzymol*, 1990, **182**, 24-38.
145. N. E. Good, G. D. Winget, W. Winter, T. N. Connolly, et al., *Biochemistry*, 1966, **5**, 467-477.
146. N. E. Good and S. Izawa, *Meth Enzymol*, 1972, **24**, 53-68.
147. W. W. Wakarchuk, A. Cunningham, D. C. Watson and N. M. Young, *Protein Eng*, 1998, **11**, 295-302.
148. J. A. Roth and M. D. Garrick, *Biochem Pharmacol*, 2003, **66**, 1-13.
149. S. L. McGovern, B. T. Helfand, B. Feng and B. K. Shoichet, *J Med Chem*, 2003, **46**, 4265-4272.
150. G. E. Tiller, T. J. Mueller, M. E. Dockter and W. G. Struve, *Anal Biochem*, 1984, **141**, 262-266.

151. J. H. Zhang, *J Biomol Screen*, 1999, **4**, 67-73.
152. L. M. Mayr and D. Bojanic, *Curr Opin Pharmacol*, 2009, **9**, 580-588.
153. G. Wu and S. K. Doberstein, *Drug Discov Today*, 2006, **11**, 718-724.
154. T. Jank, D. J. Reinert, T. Giesemann, G. E. Schulz, et al., *J Biol Chem*, 2005, **280**, 37833-37838.
155. E. Boix, G. J. Swaminathan, Y. Zhang, R. Natesh, et al., *J Biol Chem*, 2001, **276**, 48608-48614.
156. H. Jamaluddin, P. Tumbale, S. G. Withers, K. R. Acharya, et al., *J Mol Biol*, 2007, **369**, 1270-1281.
157. M. A. Haidekker and E. A. Theodorakis, *J Biol Eng*, 2010, **4**, 11.
158. M. C. Galan, A. T. Tran, K. Bromfield, S. Rabbani, et al., *Org Biomol Chem*, 2012, **10**, 7091-7097.
159. T. Ju, B. Xia, R. P. Aryal, W. Wang, et al., *Glycobiology*, 2011, **21**, 352-362.
160. X. Chen, M. J. Jou and J. Yoon, *Org Lett*, 2009, **11**, 2181-2184.
161. H. S. Lee and J. S. Thorson, *Anal Biochem*, 2011, **418**, 85-88.
162. S. Cohen, L. P. Jordheim, M. Megherbi, C. Dumontet, et al., *J Chromatogr B Analyt Technol Biomed Life Sci*, 2010, **878**, 1912-1928.
163. N. Tomiya, E. Ailor, S. M. Lawrence, M. J. Betenbaugh, et al., *Anal Biochem*, 2001, **293**, 129-137.
164. T. Ryll and R. Wagner, *J Chromatogr* 1991, **570**, 77-88.
165. J. Rabina, M. Maki, E. M. Savilahti, N. Jarvinen, et al., *Glycoconj J*, 2001, **18**, 799-805.
166. W. R. Pels Rijcken, G. J. M. Hooghwinkel and W. Ferwerda, *Biochem J*, 1990, **266**, 777-783.
167. P. Vinas, N. Campillo, I. Lopez-Garcia, S. Martinez-Lopez, et al., *J Agric Food Chem*, 2009, **57**, 7245-7249.
168. T. Cecchi, *Crit Rev Anal Chem*, 2008, **38**, 161-213.
169. N. Kochanowski, F. Blanchard, R. Cacan, F. Chirat, et al., *Anal Biochem*, 2006, **348**, 243-251.
170. K. Nakajima, S. Kitazume, T. Angata, R. Fujinawa, et al., *Glycobiology*, 2010, **20**, 865-871.
171. L. R. K. Snyder, J. J.; Dolan, J. W., *Introduction to Modern Liquid Chromatography*, A John Wiley & Sons. USA, 2010.
172. U. S. F. a. D. A. (FDA), *Center for Drug Evaluation and Research, Reviewer Guidance; Validation of Chromatographic Methods.* , Rockville, 1994.
173. A. Evitt, L. M. Tedaldi and G. K. Wagner, *Chem Commun (Camb)*, 2012, **48**, 11856-11858.
174. B. Lougheed, H. D. Ly, W. W. Wakarchuk and S. G. Withers, *J Biol Chem*, 1999, **274**, 37717-37722.
175. A. Minami, K. Kakinuma and T. Eguchi, *Tetrahedron Lett*, 2005, **46**, 6187-6190.
176. J. A. Salas and C. Mendez, *Trends Microbiol*, 2007, **15**, 219-232.
177. M. Oberthur, C. Leimkuhler, R. G. Kruger, W. Lu, et al., *J Am Chem Soc*, 2005, **127**, 10747-10752.
178. H. Jamaluddin, P. Tumbale, T. A. Ferns, N. Thiyagarajan, et al., *Biochem Biophys Res Commun*, 2009, **385**, 601-604.
179. A. Kobata, *Anal Biochem*, 1979, **100**, 1-14.
180. O. Motabar, Z. D. Shi, E. Goldin, K. Liu, et al., *Anal Biochem*, 2009, **390**, 79-84.
181. D. J. Coleman, M. J. Studler and J. J. Naleway, *Anal Biochem*, 2007, **371**, 146-153.
182. D. J. Coleman, D. A. Kuntz, M. Venkatesan, G. M. Cook, et al., *Anal Biochem*, 2010, **399**, 7-12.
183. J. M. Arrieta and G. J. Herndl, *Appl Environ Microb*, 2001, **67**, 4896-4900.

184. D. J. Urban, W. Zheng, O. Goker-Alpan, A. Jadhav, et al., *Comb Chem High Throughput Screen*, 2008, **11**, 817-824.
185. M. J. Eggertson and D. B. Craig, *Biomed Chromatogr*, 1999, **13**, 516-519.
186. A. G. Hadd, D. E. Raymond, J. W. Halliwell, S. C. Jacobson, et al., *Anal Chem*, 1997, **69**, 3407-3412.
187. E. R. Nichols, J. M. Gavina, R. G. McLeod and D. B. Craig, *Protein J*, 2007, **26**, 95-105.
188. C. Dass, *Fundamentals of Contemporary Mass Spectrometry*, John Wilet & Sons, New Jersey, 2007.
189. H. P. Nguyen and K. A. Schug, *J Sep Sci*, 2008, **31**, 1465-1480.
190. P. Jandera, *Anal Chim Acta*, 2011, **692**, 1-25.
191. T. Ikegami, K. Tomomatsu, H. Takubo, K. Horie, et al., *J Chromatogr A*, 2008, **1184**, 474-503.
192. P. Hemström and K. Irgum, *J Sep Sci*, 2006, **29**, 1784-1821.
193. E. Johnsen, S. R. Wilson, I. Odsbu, A. Krapp, et al., *J Chromatogr A*, 2011, **1218**, 5981-5986.
194. B. Preinerstorfer, S. Schiesel, M. Lammerhofer and W. Lindner, *J Chromatogr A*, 2010, **1217**, 312-328.
195. J. Nawrocki, *J Chromatogr A*, 1997, **779**, 29-71.
196. W. Naidong, *J Chromatogr B Analyt Technol Biomed Life Sci*, 2003, **796**, 209-224.
197. K. D. Wyndham, J. E. O'Gara, T. H. Walter, K. H. Glose, et al., *Anal Chem*, 2003, **75**, 6781-6788.
198. M. R. Gama, R. G. da Costa Silva, C. H. Collins and C. B. G. Bottoli, *Trends Analyt Chem*, 2012, **37**, 48-60.
199. C. Antonio, T. Larson, A. Gilday, I. Graham, et al., *Rapid Commun Mass Spectrom*, 2008, **22**, 1399-1407.
200. G. S. Philibert and S. V. Olesik, *J Chromatogr A*, 2011, **1218**, 8222-8230.
201. D. J. McNally, J. P. Hui, A. J. Aubry, K. K. Mui, et al., *J Biol Chem*, 2006, **281**, 18489-18498.
202. V. V. Tolstikov and O. Fiehn, *Anal Biochem*, 2002, **301**, 298-307.
203. Y. Guo and S. Gaiki, *J Chromatogr A*, 2005, **1074**, 71-80.
204. J. Ruta, S. Rudaz, D. V. McCalley, J. L. Veuthey, et al., *J Chromatogr A*, 2010, **1217**, 8230-8240.
205. E. de Hoffmann *J Mass Spectrom*, 1996, **31**, 129-137.
206. T. Antoine, C. Bosso, A. Heyraud and E. Samain, *Biochimie*, 2005, **87**, 197-203.
207. K. W. Harder, P. Owen, L. K. H. Wong, R. Aebersold, et al., *Biochem J*, 1994, **298**, 395-401.
208. A. R. Sherwood, B. C. Paasch, C. A. Worby and M. S. Gentry, *Anal Biochem*, 2013, **435**, 54-56.
209. B. Y. Feng and B. K. Shoichet, *Nat Protoc*, 2006, **1**, 550-553.
210. H. J. Wiggers, J. Cheleski, A. Zottis, G. Oliva, et al., *Anal Biochem*, 2007, **370**, 107-114.
211. R. Shen, S. Wang, X. Ma, J. Xian, et al., *Biochemistry (Mosc)*, 2010, **75**, 944-950.
212. J. P. Kennedy, L. Williams, T. M. Bridges, R. N. Daniels, et al., *ACS Comb Sci*, 2008, **10**.
213. C. J. Harris, R. D. Hill, D. W. Sheppard, M. J. Slater, et al., *Comb Chem High Throughput Screen*, 2011, **14**, 521-531.
214. W. R. Galloway and D. R. Spring, *Expert Opin Drug Discov*, 2009, **4**, 467-472.
215. K. A. Winans and C. R. Bertozzi, *Chem Biol*, 2002, **9**, 113-129.
216. A. L. Hopkins, C. R. Groom and A. Alex, *Drug Discov Today*, 2004, **9**, 430-431.
217. E. E. Wyatt, S. Fergus, W. R. Galloway, A. Bender, et al., *Chem Commun (Camb)*, 2006, 3296-3298.
218. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Adv Drug Delivery Rev*, 1997, **23**, 3-25.

219. G. R. Bickerton, G. V. Paolini, J. Besnard, S. Muresan, et al., *Nat Chem*, 2012, **4**, 90-98.
220. P. D. Leeson and S. A. St-Gallay, *Nat Rev Drug Discov*, 2011, **10**, 749-765.
221. A. Simeonov, A. Jadhav, C. J. Thomas, Y. Wang, et al., *J Med Chem*, 2008, **51**, 2363-2371.
222. J. Inglese, R. L. Johnson, A. Simeonov, M. Xia, et al., *Nat Chem Biol*, 2007, **3**, 466-479.
223. A. Jadhav, R. S. Ferreira, C. Klumpp, B. T. Mott, et al., *J Med Chem*, 2010, **53**, 37-51.
224. N. Thorne, D. S. Auld and J. Inglese, *Curr Opin Chem Biol*, 2010, **14**, 315-324.
225. K. E. Coan, J. Ottl and M. Klumpp, *Expert Opin Drug Discov*, 2011, **6**, 405-417.
226. A. M. Giannetti, B. D. Koch and M. F. Browner, *J Med Chem*, 2008, **51**, 574-580.
227. G. M. Rishton, *Drug Discov Today*, 2003, **8**, 86-96.
228. G. M. Rishton, *Drug Discov Today*, 1997, **2**, 382-384.
229. J. C. Hermann, Y. Chen, C. Wartchow, J. Menke, et al., *ACS Med Chem Lett*, 2013, **4**, 197-200.
230. P. A. Johnston, *Curr Opin Chem Biol*, 2011, **15**, 174-182.
231. M. Soltero-Higgin, E. E. Carlson, T. D. Gruber and L. L. Kiessling, *Nat Struct Mol Biol*, 2004, **11**, 539-543.
232. T. K. Smith, B. L. Young, H. Denton, D. L. Hughes, et al., *Bioorg Med Chem Lett*, 2009, **19**, 1749-1752.
233. T. Mendgen, C. Steuer and C. D. Klein, *J Med Chem*, 2012, **55**, 743-753.
234. E. E. Wyatt, W. R. Galloway, G. L. Thomas, M. Welch, et al., *Chem Commun (Camb)*, 2008, 4962-4964.
235. R. Milcent, J. C. Malanda, G. Barbier and J. Vaissermann, *J Heterocycl Chem*, 1997, **34**, 329-336.
236. K. E. Coan, D. A. Maltby, A. L. Burlingame and B. K. Shoichet, *J Med Chem*, 2009, **52**, 2067-2075.
237. Y. V. Frenkel, A. D. Clark, Jr., K. Das, Y. H. Wang, et al., *J Med Chem*, 2005, **48**, 1974-1983.
238. S. L. McGovern, E. Caselli, N. Grigorieff and B. K. Shoichet, *J Med Chem*, 2002, **45**, 1712-1722.
239. A. J. Ryan, N. M. Gray, P. N. Lowe and C. W. Chung, *J Med Chem*, 2003, **46**, 3448-3451.
240. K. E. Coan and B. K. Shoichet, *J Am Chem Soc*, 2008, **130**, 9606-9612.
241. B. K. Shoichet, *J Med Chem*, 2006, **49**, 7274-7277.
242. S. L. McGovern and B. K. Shoichet, *J Med Chem*, 2003, **46**, 1478-1483.
243. J. Seidler, S. L. McGovern, T. N. Doman and B. K. Shoichet, *J Med Chem*, 2003, **46**, 4477-4486.
244. L. Pohjala and P. Tammela, *Molecules*, 2012, **17**, 10774-10790.
245. B. Y. Feng, A. Shelat, T. N. Doman, R. K. Guy, et al., *Nat Chem Biol*, 2005, **1**, 146-148.
246. B. J. Berne and R. Peroca, *Dynamic Light Scattering: With Applications to Chemistry, Biology, and Physics*, Wiley, New York, 1976.
247. R. G. LeBel and A. I. Goring, *J Chem Eng Data*, 1962, **7**, 100-101.
248. J. L. Miller, *Curr Top Med Chem*, 2006, **6**, 19-29.
249. C. N. Cavasotto and S. S. Phatak, *Drug Discov Today*, 2009, **14**, 676-683.
250. F. Moreau, N. Desroy, J. M. Genevard, V. Vongsouthi, et al., *Bioorg Med Chem Lett*, 2008, **18**, 4022-4026.
251. S. Grizot, M. Salem, V. Vongsouthi, L. Durand, et al., *J Mol Biol*, 2006, **363**, 383-394.
252. R. Tripathy, A. Reiboldt, P. A. Messina, M. Iqbal, et al., *Bioorg Med Chem Lett*, 2006, **16**, 2158-2162.
253. R. Tripathy, A. Ghose, J. Singh, E. R. Bacon, et al., *Bioorg Med Chem Lett*, 2007, **17**, 1793-1798.
254. V. Hadi, Y. H. Koh, T. W. Sanchez, D. Barrios, et al., *Bioorg Med Chem Lett*, 2010, **20**, 6854-6857.

255. A. Dandia and A. K. Jain, *J Heterocyclic Chem*, 2013, **50**, 104-113.
256. N. M. Parekh and K. C. Maheria, *Med Chemi Res*, 2012, **21**, 4168-4176.
257. A. Cornish-Bowden, *Fundamentals of enzyme kinetics* Portland Press, London, 2004.
258. A. Cornish-Bowden, *Biochem J*, 1974, **137**, 143-144.
259. M. Levy, presented in part at the International Symposium on Aspirin Intolerance and Related Syndromes, Rome, Italy, 2000.
260. I. Nikolova, J. Tencheva, J. Voinikov, V. Petkova, et al., *Biotechnol Biotec Eq*, 2012, **26**, 3329-3337.
261. H. Lineweaver and D. Burk, *J Am Chem Soc*, 1934, **56**, 658-666.
262. M. Dixon, *Biochem J*, 1953, **55**, 170-171.
263. W. A., *Enzymes in Industry: Production and Applications*, WILEY-VCH, 2007.